COURTESY COPY OF THE PCT APPLICATION AS ORIGINALLY FILED WITH SEQUENCE LISTING AND ABSTRACT

A method of modulating cell survival, differentiation and/or synaptic plasticity

Field of invention

The present invention relates to a method of modulating cell differentiation and/or survival by providing compounds capable of modulating the interaction between two individual neural cell adhesion molecules (NCAM). The invention further relates to a method of screening for candidate compounds capable of modulating the interaction between the Ig1, Ig2 and/or Ig3 modules of two individual NCAM. The invention also concerns the use of the identified candidate compounds for the manufacture of a medicament.

Background of invention

The neural cell adhesion molecule, NCAM, mediates cell-cell adhesion via homophilic (NCAM-NCAM) binding. NCAM plays a key role in neural development, neuronal differentiation and synaptic plasticity, including learning and memory consolidation.

Intercellular interactions play a crucial role in a wide range of biological processes, including cell migration, survival and differentiation. These phenomena depend upon protein recognition at the cell surface mediated by cell-cell adhesion molecules (CAMs).

The neural cell adhesion molecule, NCAM, originally described as a synaptic membrane protein (Jørgensen and Bock, 1974), and later shown to mediate cell-cell adhesion was the first mammalian cell adhesion molecule identified. NCAM belongs to the immunoglobulin (Ig) superfamily. Alternative splicing of mRNA and post-translational modifications generate a large number of NCAM isoforms. The three major NCAM isoforms have identical extracellular parts consisting of five Ig modules and two fibronectin type III modules.

NCAM is known to mediate Ca^{2+}-independent cell-cell and cell-substratum adhesion via homophilic (NCAM binding to NCAM) and heterophilic (NCAM binding to other molecules) interactions (Berezin et al., 2000). The different modules of NCAM have

been shown to perform distinct functions. NCAM binds various extracellular matrix components such as heparin/heparan sulfate, chondroitin sulfate proteoglycans, and different types of collagen. The heparin binding sequence is localized to the Ig2 module. NCAM also binds to the neural cell adhesion molecule L1. This interaction is believed to take place between the fourth Ig module of NCAM and an oligomannosidic moiety expressed on L1.

Despite extensive studies, the precise mechanism of the homophilic binding of NCAM remains unclear, and the published results are to some extent contradictory. NCAM homophilic binding was originally reported to depend on an antiparallel interaction between Ig3 modules from two opposing NCAM molecules. Cell aggregation experiments performed on mouse L-cells expressing chicken NCAM with deletions of different Ig modules indicated an involvement of the Ig3 module. Later, employing microspheres coated with individual recombinant Ig modules of chicken NCAM, binding was demonstrated between the Ig1 and Ig5 modules, and between the Ig2 and Ig4 modules, whereas microspheres coated with Ig3 exhibited strong self-aggregation (Ranheim et al., 1996). However, a study by Atkins et al. (2001) on the solution structure of the Ig3 module of chicken NCAM including ultracentrifugation experiments did not support the suggested dimerization of Ig3.

A binding between recombinant modules of rat Ig1 and Ig2 was demonstrated by means of surface plasmon resonance analysis (Kiselyov et al., 1997). The three-dimensional structures of individual modules of rat Ig1 and Ig2, and the chicken Ig1 module, have been determined by nuclear magnetic resonance (NMR) spectroscopy, resulting in the identification of amino acid residues involved in the homophilic binding between the Ig1 and Ig2 modules (Thomsen et al., 1996; Jensen et al., 1999; Atkins et al., 1999). The crystal structure of the Ig1-2 fragment of rat NCAM provided detailed information on the cross-like Ig1-2 dimer, and pointed out the key residues in this interaction, namely F19 and Y65 (Kasper et al., 2000). Recently, it was demonstrated that a point mutation of F19 (F19S) did not affect cell aggregation mediated by full length NCAM, even though it abolished dimerization of the Ig1-2-3 fragment, which otherwise takes place in solution (Atkins et al., 2001). These results therefore question the suggested Ig3-to-Ig3 (Rao et al., 1992; Ranheim et al., 1996) and Ig1-to-Ig2 (Kiselyov et al., 1997; Kasper et al., 2000) models of NCAM homophilic binding.

Thus, two non-overlapping homophylic binding sites of NCAM have been described in scientific literature: the Ig3-to-Ig3 and Ig1-to-Ig2 binding sites. The sequences derived from these two sites have been shown to be capable of stimulating neurite outgrowth and modulating adhesion of neural cells (WO03020749, Soroka et al, 2002; Rao et al., 1992; Ranheim et al., 1996). It has also been shown that peptide sequences, which are capable to bind to the Ig3-to-Ig3 binding site, do not interfere with the biological effects mediated by the Ig1-to-Ig2 binding site, and vice versa. The latter finding indicate that NCAM homophylic adhesion has a much more complex mechanism, then just the mechanism of mechanistic binding of two individual NCAM molecules though the multiple homophylic binding sites, and that the involvement of one or another homophylic binding site in a process mediated by NCAM may depend on a particular NCAM environment, such as for example the presence of a ligand of one or another binding site, or availability of one or another site for binding.

The present invention provides a method of modulating such processes by providing compounds capable of binding to NCAM modules Ig1, Ig2 and/or Ig3 through a novel homophylic binding site.

Summary of invention

According to the present invention, providing the ligands directed to different homophylic binding sites of NCAM may allow a fine regulation of involvement of NCAM in molecular mechanisms underlying different processes related to placticity of neural cells and thereby modulation of these processes. Thus, the present invention concerns the compounds, which are capable of modulating adhesion, induce differentiation, and promote regeneration, neuronal plasticity and survival of cells expressing NCAM, and methods for screening and using such compounds.

In one aspect the present invention relates to a method of modulating cell differentiation and/or survival of the neural cell adhesion molecule (NCAM) presenting cells comprising

a) providing a candidate compound capable of

i) interacting with the Ig1 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig1 and Ig3 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

ii) interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig1 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

iii) interacting with the Ig2 module of NCAM, and thereby mimicking the interaction between Ig2 and Ig3 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

iv) interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig2 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

v) interacting with the Ig2 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig2 and Ig2 modules of NCAM, wherein said modules are from two individual NCAM molecules,

b) providing at least one NCAM presenting cell;

c) contacting the at least one NCAM presenting cell with said candidate compound, and thereby modulating cell differentiation and/or survival of the at least one NCAM presenting cell.

In another aspect the present invention is concerned with a method for testing a compound whether it is capable of modulating interaction between two individual NCAM molecules through a homophylic binding site composed of the Ig1, Ig2 and Ig3 modules of said NCAM molecules by modulating the interaction of

i) the Ig1 module of one individual NCAM molecule with the Ig3 module of another individual NCAM molecule, and/or

ii) the Ig2 module of one individual NCAM molecule with the Ig3 module of another individual NCAM molecule, and/or

iii) the Ig2 module of one individual NCAM molecule with the Ig2 module of another individual NCAM molecule

said method comprising

a) providing a compound;

b) providing at least one individual fragment of an NCAM molecule, wherein said fragment comprising a sequence of consecutive amino acid residues corresponding to the sequence of the Ig1-2-3 module of NCAM comprising

residues 1 to 289 of the sequence set forth in SEQ ID NO: 44, or a fragment of said individual fragment;

c) testing whether the compound is capable of

i) interacting with the Ig1 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig1 and Ig3 modules of NCAM, wherein said modules are from the two individual fragments of (b) interacting to each other, and/or

ii) interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig1 modules of NCAM, wherein said modules are from the two individual fragments of (b) interacting to each other, and/or

iii) interacting with the Ig2 module of NCAM, and thereby mimicking the interaction between Ig2 and Ig3 modules of NCAM, wherein said modules are from the two individual fragments of (b) interacting to each other, and/or

iv) interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig2 modules of NCAM, wherein said modules are from the two individual fragments of (b) interacting to each other, and/or

v) interacting with the Ig2 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig2 and Ig2 modules of NCAM, wherein said modules are from the two individual fragments of (b) interacting to each other.

by contacting the compound with a fragment of (b);

d) selecting a compound capable of at least one interaction of (c) as a candidate compound capable of modulating differentiation, adhesion and/or survival of a cell presenting NCAM.

In still another aspect the present invention provide a method for selecting a candidate compound capable of modulating differentiation, adhesion and/or survival of NCAM presenting cells by modulating the interaction of

i) the Ig1 module of one individual NCAM molecule with the Ig3 module of another individual NCAM molecule, and/or

ii) the Ig2 module of one individual NCAM molecule with the Ig3 module of another individual NCAM molecule, and/or

iii) the Ig2 module of one individual NCAM molecule with the Ig2 module of another individual NCAM molecule,

said method comprising the steps of

a) providing a soluble or a crystalline polypeptide comprising the Ig1-2-3 module of NCAM,

b) generating a structural model of the Ig1-2-3 module of NCAM of (a) by using the computer modelling techniques;

c) in-silico evaluating compounds for the capability of

i) interacting with the Ig1 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig1 and Ig3 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

ii) interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig1 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

iii) interacting with the Ig2 module of NCAM, and thereby mimicking the interaction between Ig2 and Ig3 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

iv) interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig2 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

v) interacting with the Ig2 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig2 and Ig2 modules of NCAM, wherein said modules are from two individual NCAM molecules,

by using the structural model(s) of the Ig1-2-3 module of NCAM of (c);

d) selecting a candidate compound capable of at least one interaction of (b), and

e) testing the candidate compound of (d) in an in vitro or in vivo assay for the capability of modulating differentiation, adhesion and/or survival of NCAM presenting cells, said assays comprising at least one NCAM presenting cell, and/or

f) testing the candidate compound of (d) in an assay comprising evaluating the capability of the compound of at least one interaction of (b) by contacting the compound with at least one individual fragment of an NCAM molecule, said fragment comprising a sequence of consecutive amino acid residues

corresponding to the sequence of the Ig1-2-3 module of NCAM comprising residues 1 to 289 of the sequence set forth in SEQ ID NO: 44.

Thus, it is an objective of the present invention to provide a crystalline protein comprising the Ig1-2-3 module of NCAM and a method of preparing said crystalline protein.

The invention further provides the compounds capable of

i) interacting with the Ig1 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig1 and Ig3 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

ii) interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig1 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

iii) interacting with the Ig2 module of NCAM, and thereby mimicking the interaction between Ig2 and Ig3 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

iv) interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig2 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

v) interacting with the Ig2 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig2 and Ig2 modules of NCAM, wherein said modules are from two individual NCAM molecules.

The invention also concerns using the compounds selected by the above methods for the manufacture of a medicament and features pharmaceutical compositions comprising thereof.

Description of Drawings

Figure 1 presents crystallographic data and refinement statistics.

Figure 2 presents the atomic structure coordinates of the Ig1-2-3 module crystal.

Figure 3. Crystal structure of the rat NCAM Ig1-2-3 fragment at 2.0 Å resolution. **(A)** Cα backbone diagram in stereo with every 10^{th} residue labeled.

(B) Ribbon diagram with β-strands labeled according to Ig I set nomenclature.

Figure 4. Crystal structure of the Ig1-2-3 fragment of NCAM reveals four major module-module interactions and two kinds of Ig1-2-3 arrays. Space-filling models of interacting Ig1-2-3 cis dimers (mediated by Ig1-Ig2 binding) are shown. The Ig1-to-Ig2, Ig1-to-Ig3, Ig2-to-Ig2, and Ig2-to-Ig3 interaction sites are indicated by white ellipses. The heparin binding sites of the Ig2 modules (residues 133-148) are indicated by white squares. The arrows indicate the position of N-linked glycosylation at Asn203 (Asn203 is marked by white triangle). The termini are denoted by N and C.

(A,B) The Ig1-2 mediated cis dimers of the Ig1-2-3 fragment form a "flat" zipper via an Ig2-to-Ig3 mediated trans interaction, reflecting an interaction between NCAM molecules on opposing cells.

(C,D) The Ig1-2-3 fragment cis dimers also form a non-symmetrical "compact" zipper via Ig1-to-Ig3 and Ig2-to-Ig2 trans interactions. Two cis dimers are held together by two Ig1-to-Ig3 interactions (full ellipses) on one side and one Ig2-to-Ig2 interaction (stippled ellipse) on the opposite side of the zipper. The views in **B** and **D** are perpendicular to **A** and **C**, respectively.

Figure 5. Close-up view of the interaction interfaces in the NCAM Ig1-2-3 fragment.

(A) The Ig1-to-Ig2 interaction interface. The Ig1 and Ig2 modules are belong to two different individual Ig1-2-3 fragments that form one Ig1-2-3 cis dimer.

(B) The Ig2-to-Ig3 interaction interface.

(C) The Ig2-to-Ig2 interaction interface.

(D) The Ig1-to-Ig3 interaction interface. In **B-D**, the ribbon representations of modules from two interacting Ig1-2-3 fragments belonging to different individual Ig1-2-3 cis dimers. The hydrogen bonds are shown as dashed lines.

Figure 6. The effect of the Ig3 module, the P1-B, P3-DE, P3-G, P3-B peptides, and their derivatives, on neurite outgrowth from the NCAM-expressing PC12-E2 cells grown on top of a confluent monolayer of NCAM-transfected fibroblasts.

(A-F) Confocal micrographs of NCAM-expressing pheochromocytoma PC12-E2 cells grown on top of a confluent monolayer of vector-transfected **A,C,E** or NCAM-140 transfected **B,D,F** L929 fibroblasts. NCAM-NCAM interaction stimulates neurite outgrowth in PC12-E2 cells grown on top of NCAM-expressing (LBN) **B** versus NCAM-

negative (LVN) **A** fibroblasts. Introduction of the recombinant Ig3 module does not affect PC12-E2 cells grown on vector-transfected fibroblasts **C** but clearly inhibits neurite outgrowth in PC12-E2 cells grown on NCAM-transfected fibroblasts **D** as a result of disruption of NCAM-NCAM interactions. In contrast, Ig3mut2 neither affects PC12-E2 cells grown on vector-transfected fibroblasts **E** nor inhibits NCAM-induced neurite outgrowth **F**. Peptides P1-B, P3-DE, and P3-G have inhibitory effects comparable to the effect of Ig3wt **C,D**, whereas effects of the Ig3mut1, P3-B peptide, and control peptides are similar to the effect of Ig3mut2 **E,F**. Scale bar, 20 μm.

(G) The effect of the Ig3 module, P1-B, P3-DE, P3-G, P3-B peptides, and their derivatives, is shown as percent of control, setting the difference between the average neurite length of PC12-E2 cells grown on NCAM-140-transfected and vector-transfected fibroblasts to 100%. Results are given as mean ± SEM.* $P<0.05$, ** $P<0.01$ (compared to the induction of neurite outgrowth from PC12-E2 cells grown on top of monolayer of NCAM-transfected fibroblasts).

Figure 7. Schematic representations of the "compact", "flat", and "double" zipper adhesion complexes formed by NCAM, as observed in the crystal structure of the NCAM Ig1-2-3 fragment. The individual NCAM modules are shown as cylinders. The Ig and FnIII modules are numbered by Arabic and Roman numerals, respectively. In order to accommodate all seven extracellular modules of NCAM a bend has been introduced after Ig4 according to electron microscopy studies (Hall and Rutishauser, 1987; Becker et al., 1989). The size of the Ig1-2-3 fragment and distance between opposing cell membranes are indicated.

(A) The "compact" zippers are stabilized by Ig1-to-Ig3 and Ig2-to-Ig2 interactions between Ig1-2-3 cis dimers originating from two opposing cell membranes.

(B) The "flat" zipper is stabilized by Ig2-to-Ig3 interactions between Ig1-2-3 cis dimers originating from two opposing cell membranes.

(C) The two types of zippers may co-exist as observed in the crystal and will result in formation of a double zipper-like adhesion complex.

Figure 8. demonstrates the effect of the P2-CD peptide on neurite outgrowth of CGN grown as single neurons in primary culture for 24 h in the presence of different concentrations of the peptide in growth media. The length of neurites is expressed in Arbitrary Units (AU). The length of neurites in treated cultures is compared to the length of neurite in cultures without treatment (control) (* $p<0,05$; **$p<0,02$;

***p<0,001 ****p<0,0005). P2d, which is a peptide fragment of NCAM Ig2 module (see Soroka et al., 2002), was used as a positive control to indicate responsiveness of the cells to the treatment.

Figure 9 demonstrates the effect of the P3-G peptide on neurite outgrowth of CGN grown as single neurons in primary culture for 24 h in the presence of different concentrations of the peptide in growth media. The length of neurites is expressed in Arbitrary Units (AU). The length of neurites in treated cultures is compared to the length of neurite in cultures without treatment (control). P2d was used as a positive control. * p<0,05; **p<0,02; ***p<0,001 ****p<0,0005

Figure 10 demonstrates the effect of the P2-EF peptide on neurite outgrowth of CGN grown as single neurons in primary culture for 24 h in the presence of different concentrations of the peptide in growth media. The length of neurites is expressed in Arbitrary Units (AU). The length of neurites in treated cultures is compared to the length of neurite in cultures without treatment (control). P2d as used as a positive control. * p<0,05; **p<0,02; ***p<0,001 ****p<0,0005

Figure 11 demonstrates the effects of the P1-CD peptide on neurite outgrowth of CGN grown as single neurons in primary culture for 24 h in the presence of different concentrations of the peptide in growth media and in co-culture of CGN with genetically modified fibroblasts with (LBN) or without (LVN) NCAM expression. It can be seen that the peptide does not affect NCAM-independent neurite outgrowth of CGN in both cultures, but inhibits NCAM-dependent neurite outgrowth by interfering with NCAM homophylic adhesion in co-cultures of CGN and LBN-cells. The length of neurites is expressed in Arbitrary Units (AU). The length of neurites in treated cultures is compared to the length of neurite in cultures without treatment (control). P2d was used as a positive control. ***p<0,001

Figure 12 demonstrates the effect of the P2-A'B peptide on neurite outgrowth of CGN grown as single neurons in primary culture for 24 h in the presence of different concentrations of the peptide in growth media and in co-culture of CGN with genetically modified fibroblasts with (LBN) or without (LVN) NCAM expression. From the figure it can be seen that the peptide does not affect NCAM-independent neurite outgrowth of CGN in both cultures, and inhibits NCAM-dependent neurite outgrowth

by interfering with NCAM homophylic adhesion in co-cultures of CGN and LBN-cells. The length of neurites is expressed in Arbitrary Units (AU). The length of neurites in treated cultures is compared to the length of neurite in cultures without treatment (control). P2d was used as a positive control. ***p<0,001

Detailed description of the invention

Molecules with the potential to promote neurite outgrowth as well as stimulate survival, regeneration and modulate adhesion of neuronal cells, such as certain endogenous trophic factors and adhesion molecules, for example NCAM, are prime targets in the search for compounds that facilitate for example neuronal regeneration and other forms of neuronal plasticity. To evaluate the potential of compounds to interfere with cell adhesion, the ability to stimulate neurite outgrowth, regeneration and the survival of neuronal cells a capability of the compounds to interact with NCAM may be investigated. It is an object of the present invention to provide compounds capable of binding to one or more positions in the NCAM molecule. In particular, positions in the NCAM Ig1, Ig2 and/or Ig3 modules which constitute a homophylic binding site of NCAM described in the present application.

NCAM is a multifunctional adhesion molecule. It is involved as a key molecule in different processes associated with neural plasticity during embryonic development, in the adult brain and in association with disease. Involvement of NCAM in different processes underlying neural plasticity is provided by a capability of NCAM to cis- and trans-homophylic interactions and heterophylic interactions with a number of cellular receptors and other cellular and extracellular molecules. The NCAM molecule has multiple non-ovelapping and overlapping binding sites for the interaction with these molecules, which are located in different extracellulr NCAM modules and in the intracellular domain of NCAM.

The present invention relates to a method of modulating the differentiation, adhesion and/or survival of NCAM presenting cells, said method comprising providing a compound capable of interacting with a novel NCAM homophylic binding site composed of amino acid residues of the Ig1, Ig2 and Ig3 modules of NCAM. Amino acid residues of the binding site are capable of the following interactions:

- amino acid residues of the binding site located in the Ig1 module of one

NCAM molecule of the interaction with the amino acid residues of the binding site located in the Ig3 module of another, the counter NCAM molecule, but not with the residues of the binding site located in Ig1 or Ig2 modules of this counter NCAM molecule,

- amino acid residues of the binding site located in the Ig2 module of one NCAM molecule of interaction with the amino acid residues of the binding site located in the Ig2 module of another, the counter NCAM molecule, but not with the residues of the binding site located in the Ig1 or Ig3 modules of this counter NCAM molecule, and
- amino acid residues of the binding site located in the Ig2 module of one NCAM molecule of the interaction with the amino acid residues of the binding site located in the Ig3 module of another, the counter NCAM molecule, but not with the residues of the binding site located in the Ig2 or Ig1 modules of this counter NCAM molecule,
- amino acid residues of the binding site located in the Ig3 module of one NCAM molecule of the interaction with the amino acid residues of the binding site located in the Ig2 module of another, the counter NCAM molecule, but not with the residues of the binding site located in the Ig3 or Ig1 modules of this counter NCAM molecule,
- amino acid residues of the binding site located in the Ig3 module of one NCAM molecule of the interaction with the amino acid residues of the binding site located on the Ig1 module of another, counter NCAM molecule, but not with the residues of the binding site located on the Ig3 or Ig2 modules of this counter NCAM molecule.

According to the invention, a compound, which is capable of

i) interacting with the Ig1 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig1 and Ig3 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

ii) interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig1 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

iii) interacting with the Ig2 module of NCAM, and thereby mimicking the interaction between Ig2 and Ig3 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

iv) interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig2 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

v) interacting with the Ig2 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig2 and Ig2 modules of NCAM, wherein said modules are from two individual NCAM molecules,

is a ligand of the above homophylic binding site and is capable of modulating a process assisted by NCAM homophylic binding through the above binding site by interacting with this binding site.

The term "individual" in relation to two or more molecules/modules/fragments is used to indicate that these two or more molecules/modules/fragments are present as separate, non-connected, substances.

The term "interacting" is used herein synonymously with the term "binding". The term "ligand" is defined as a compound, which binds to the binding site of above and mimics NCAM homophylic binding. The term "to mimic" is understood as a capability of the ligand to induce/stimulate or inhibit a biological process, which is mediated by NCAM through homophylic interaction through the above binding site. The ligand may also inhibit naturally occurring interactions, such as by binding to parts of NCAM which are not a part of the binding site, and wherein the interference is merely a steric interference.

The compounds capable of interaction/binding to the binding site of the invention are favourable for the promotion of neurite outgrowth. Compounds of the invention are therefore considered to be good promoters of regeneration of neuronal connections, and thereby of functional recovery after damages, as well as promoters of neuronal function in other conditions where such an effect is required.

In the present context "differentiation" is related to the processes of maturation of cells, such as for example extension of neurites from neurons which takes place after the last cell division of said neurons has ended. The compounds of the present invention may be capable of stopping cell division and initiate maturation and/or extension of neurites

In the present invention a compound is considered promising when it is capable of stimulating neurite outgrowth, for example when it is capable of stimulating neurite outgrowth of cultured cells when compared to control cells, such as improving neurite outgrowth by 50% or more, such as 75%, for example 100 % or more.

Further, in the present context the wording "stimulate/promoting survival" is used synonymously with the wording "preventing cell death" or "neuro-protection". By stimulating/promoting survival it is possible to prevent diseases or prevent further degeneration of the nervous system in individuals suffering from a neuro-degenerative disorder.

"Survival" refers to the process, wherein a cell has been traumatised and would under normal circumstances, with a high probability die, if not the compound of the invention was used to prevent said cell from degenerating, and thus promoting or stimulating survival of said traumatised cell.

By the term "modulation" is meant a change, such as either stimulation or inhibition. A compound of the invention is capable of modulation of the processes mediated by NCAM homophylic binding. Thus, the compound is capable of stimulation or inhibition of neural cell differentiation and/or survival. The latter processes may involve activating or inhibiting of NCAM signalling. Thus, the compound, which is capable of modulating the above NCAM dependent processes, is considered by the invention as a compound, which is capable of modulating NCAM signalling. Under the capability of a compound "to modulate the NCAM signalling" is understood a capability of a molecule to modulate the process of initiating of the production of second messenger molecules and/or activation or inhibition of an intracellular cascade reaction leading to a physiological response of the cell, such as for example an increase in neurite length in response to ligand binding to the homophylic binding site of the invention.

The invention also provides for a compound capable of "interfering with cell adhesion". This refers to the process wherein cells are attracted to one another and where the present compound is capable of either stimulating or inhibiting said attraction.

The compounds according to the invention also relates to the prevention of neuronal cell death. Peripheral nerve cells possess to a limited extent a potential to regenerate and re-establish functional connections with their targets after various injuries. However, functional recovery is rarely complete and peripheral nerve cell damage remains a considerable problem. In the central nervous system, the potential for regeneration is even more limited. Therefore, the identification of substances with the ability to prevent neuronal cell death in the peripheral and the central nervous system is significant and of great commercial value.

The compounds of the invention may be peptides, such as the peptide fragments/parts of the binding site, or peptides comprising the amino acid sequences of the binding site.

In a further embodiment of the invention the compounds may comprise other chemical entities, such as sugar, cholesterol, and fatty acid. Preferably, the chemical entity is bound to the N-terminal or C-terminal of the peptide of the compound.

It is an aspect of the present invention that the compounds are capable of binding to the NCAM Ig1 and/or Ig2 and/or Ig3 modules at the homophilic binding site of the invention, or at any other sites of the NCAM module consisting of the Ig1, Ig2 and Ig3 modules and mimicking the effect of the binding at said homophilic binding site, or modulating said effect.

Without being bound by theory, the present inventors believe that active ligands to the NCAM Ig1 and/or Ig2 and/or Ig3 modules are ligands which bind to the NCAM Ig1 and/or Ig2 and/or Ig3 modules and thus trigger a conformational change of the module resulting in a signalling cascade being initiated, wherein said signalling results in a physiological change in the cell, such as influencing survival of cells, cellular adhesion and/or neurite outgrowth. Thus, a compound according to the invention may be any compound described above which can trigger a conformational change of the NCAM Ig1 and/or the NCAM Ig2 and/or the NCAM Ig3 module resulting in a change in downstream signalling cascade.

Method of modulating

Thus, it is an object of the present invention to provide a method of modulating adhesion, differentiation and/or survival of NCAM presenting cells by

a) providing a compound capable of interacting with an NCAM homophylic binding site composed of the amino acid residues of the Ig1, Ig2 and Ig3 modules of NCAM by

i) interacting with the Ig1 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig1 and Ig3 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

ii) interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig1 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

iii) interacting with the Ig2 module of NCAM, and thereby mimicking the interaction between Ig2 and Ig3 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

iv) interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig2 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

v) interacting with the Ig2 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig2 and Ig2 modules of NCAM, wherein said modules are from two individual NCAM molecules,

b) providing at least one NCAM presenting cell;

c) modulating cell differentiation and/or survival of the at least one NCAM presenting cell by contacting the at least one NCAM presenting cell with said compound, and thereby .

The invention concerns the NCAM presenting cell being

i) a cell, which naturally express an NCAM molecule on the cell surface, such as for example a neural cell, a muscle cell or a cell of any other tissue,

ii) a cancer cell, which express an NCAM molecule on the cell surface;

iii) a recombinant cell, which was genetically modified to express an NCAM molecule on the cell surface.

The NCAM presenting cell of above may be an in vivo cell, such as a cell of the body of an animal, or it may be a cell cultured in vitro. Accordingly, the above

method may be used for modulating differentiation, survival and/or adhesion of NCAM presenting cells both in vivo and in vitro. In some embodiments the method is for the in vitro use, in other embodiments the method is for the in vivo use.

The method of above comprises providing a compound capable of interacting with the NCAM homophylic binding site composed by the amino acid residues of the Ig1, Ig2 and Ig3 modules of NCAM. In one embodiment it may be a compound capable of interacting with the residues of Ig1 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig1 and Ig3 modules of NCAM, wherein said modules are from two individual NCAM molecules.

In another embodiment this may be a compound capable of interacting with the residues of Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig1 modules of NCAM, wherein said modules are from two individual NCAM molecules.

In still another embodiment the compound is capable of interacting with the residues of the Ig2 module, and thereby mimicking the interaction between Ig2 and Ig3 modules of NCAM, wherein said modules are from two individual NCAM molecules.

In yet another embodiment the compound may be capable of interacting with the residues of the Ig3 module, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig2 modules of NCAM, wherein said modules are from two individual NCAM molecules.

In yet still another embodiment the compound may be capable of interacting with the residues of the Ig2 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig2 and Ig2 modules of NCAM, wherein said modules are from two individual NCAM molecules.

The above compounds may be represented by

a) the molecules, which are capable of all the above interactions (i) to (v), or some of the above interaction, such as for example the interaction of (i) and (ii), or (i) and (iii) or any other combinations of the interactions of (i) to (v), or

b) the molecules, which are capable of only one interaction selected from any of the interactions (i) to (v).

Providing a compound capable of one or more of the above interactions may be done in one embodiment by using a method for selecting a candidate compound described in the present application. In another embodiment it may be dome by using a method for testing a compound described in the application below.

All the compounds of (a) and (b) are presumed to be capable of modulating the NCAM functions, if the executing of said functions by an NCAM molecule involves one or more interactions (i) to (v) of above.

Under the compound capable of "mimicking" the interaction means a compound acting as a ligand of the homophylic binding site of above capable of binding to the Ig1, Ig2 or Ig3 modules, and thereby replacing the binding to these modules of the Ig3, Ig2 or Ig1 modules of another, the counter, NCAM molecule, respectfully, as described above. The mimicking is results according to the invention in stimulating or inhibiting a biological process related to the latter binding.

The present inventors present herein a model for NCAM homophilic binding, wherein the Ig1 and Ig2 modules mediate dimerization of individual NCAM molecules situated on the same cell surface (*cis* interaction), and wherein the Ig3 module mediates interactions between individual NCAM molecules expressed on the surface of opposing cells (*trans* interaction) through simultaneous binding to the Ig1 and Ig2 modules. This arrangement results in the formation of a double zipper-like NCAM adhesion complex.

Sequences from NCAM

A compound of the invention may be a peptide fragment derived from the sequence of NCAM, or a variant of said peptide fragment. The peptide fragment may be a fragment of the NCAM sequence identified as SwissProt accession number NP_113709 (SEQ ID NO: 44) or SwissProt accession number P13591 (SEQ ID NO: 45).

A preferred peptide fragment may be selected from the amino acid sequences identified below:

WFSPNGEKLSPNQ (SEQ ID NO: 1)
YKCVVTAEDGTQSE (SEQ ID NO: 2)
TLVADADGFPEP (SEQ ID NO: 3)
QIRGIKKTD (SEQ ID NO: 4)
DVR (SEQ ID NO: 5)
RGIKKTD (SEQ ID NO: 6)
DVRRGIKKTD (SEQ ID NO: 7
KEGED (SEQ ID NO: 8)
IRGIKKTD (SEQ ID NO: 9)
KEGEDGIRGIKKTD (SEQ ID NO: 10)
DKNDE (SEQ ID NO: 11)
TVQARNSIVNAT (SEQ ID NO: 12)
SIHLKVFAK (SEQ ID NO: 13)
LSNNYLQIR (SEQ ID NO: 14)
RFIVLSNNYLQI (SEQ ID NO: 15)
KKDVRFIVLSNNYLQI (SEQ ID NO: 16)
QEFKEGEDAVIV (SEQ ID NO: 17)
KEGEDAVIVCD (SEQ ID NO: 18)
GEISVGESKFFL (SEQ ID NO: 19)
KHIFSDDSSELTIRNVDKNDE (SEQ ID NO: 20),
AFSPNGEKLSPNQ (SEQ ID NO: 40),
AKSVVTAEDGTQSE (SEQ ID NO: 41)
DVRRGIKKTD (SEQ ID NO: 42)
QIRGIKKTD (SEQ ID NO: 43).

The above amino acid sequences are derived from the sequence of rat NCAM having the SwissProt accession number NP_113709 (SEQ ID NO: 40).

Another preferred peptide fragment may be selected from fragments or variants of the above identified sequences.

The "variant" is to be understood as being any peptide sequence capable of interacting with the Ig1, Ig2 and/or Ig3 modules of NCAM, and via said interacting

induce differentiation, modulate cellular adhesion, stimulate regeneration, neuronal plasticity and survival of cells. Thus, fragment or variant is a biologically active compound and may be defined as a compound

i) comprising an amino acid sequence capable of being recognised by an antibody , which is also capable of recognising the predetermined NCAM amino acid sequence, and/or

ii) comprising an amino acid sequence capable of binding to a receptor moiety, which is also capable of binding the predetermined NCAM amino acid sequence, and/or

iii) having a substantially similar binding affinity to at least one of the Ig1, Ig2 or Ig3 modules as said predetermined NCAM amino acid sequence.

Thus, according to the invention, a variant as defined above is the functional equivalent of a preferred peptide fragment of above.

A variant of the full length NCAM protein, such as the NCAM of SEQ ID NO: 44 or 45, may be represented by a natural isoform of the protein, such as natural soluble molecules of NCAM, shorter or longer polypeptides of NCAM generated as a result of alternative splicing, or it may be a recombinant protein containing a fragment of NCAM comprising 30-100% of the residues the full length protein, or it may a natural protein, which has homology to NCAM. The homology between amino acid sequences may be calculated using well known algorithms such as BLOSUM 30, BLOSUM 40, BLOSUM 45, BLOSUM 50, BLOSUM 55, BLOSUM 60, BLOSUM 62, BLOSUM 65, BLOSUM 70, BLOSUM 75, BLOSUM 80, BLOSUM 85, or BLOSUM 90. Homologues are to be considered as falling within the scope of the present invention when they are at least about 40 percent homologous with the NCAM of SEQ ID NO: 44 or 45, such as at least about 50 percent homologous, for example at least about 60 percent homologous, such as at least about 70 percent homologous, for example at least about 75 percent homologous, such as at least about 80 percent homologous, for example at least about 85 percent homologous, such as at least about 90 percent homologous, for example at least 92 percent homologous, such as at least 94 percent homologous, for example at least 95 percent homologous, such as at least 96 percent homologous, for example at least 97 percent homologous, such as at least 98 percent homologous, for example at least

99 percent homologous. According to one embodiment of the invention the homology percentages refer to identity percentages.

The NCAM variant is according to invention a functional variant, such as an variant that remains a capability of the full length protein to homophylic binding through the binding site of the invention and executing the functions assisted by this binding.

The binding affinity of the compound according to the invention preferably has a binding affinity (Kd value) to the NCAM modules in the range of 10^{-3} to 10^{-10} M, such as preferably in the range of 10^{-4} to 10^{-8} M. According to the present invention the binding affinity is determined by one of the following assays of surface plasmon resonance analysis or nuclear magnetic resonance spectroscopy.

In one embodiment variants may be understood as exhibiting amino acid sequences gradually differing from the preferred predetermined sequence, as the number and scope of insertions, deletions and substitutions including conservative substitutions increase. This difference is measured as a reduction in homology between the predetermined sequence and the variant.

"Variants of peptide sequences" means that the peptides may be modified, for example by substitution of one or more of the amino acid residues. Both L-amino acids and D-amino acids may be used. Other modification may comprise derivatives such as esters, sugars, etc. Examples are methyl and acetyl esters. Polymerisation such as repetitive sequences or attachment to various carriers are well-known in the art, e.g. lysine backbones, such as lysine dendrimers carrying 4 peptides, 8 peptides, 16 peptides, or 32 peptides. Other carriers may be protein moieties, such as bovine serum albumin (BSA), or lipophilic dendrimers, or micelle-like carriers formed by lipophilic derivatives, or starburst (star-like) carbon chain polymer conjugates, or ligand presenting assembly (LPA) based on derivatives of diethylaminomethane.

Variants of the peptide fragments according to the invention may comprise, within the same variant, or fragments thereof or among different variants, or fragments thereof, at least one substitution, such as a plurality of substitutions introduced independently of one another. Variants of the complex, or fragments thereof may

thus comprise conservative substitutions independently of one another, wherein at least one glycine (Gly) of said variant, or fragments thereof is substituted with an amino acid selected from the group of amino acids consisting of Ala, Val, Leu, and Ile, and independently thereof, variants, or fragments thereof, wherein at least one alanine (Ala) of said variants, or fragments thereof is substituted with an amino acid selected from the group of amino acids consisting of Gly, Val, Leu, and Ile, and independently thereof, variants, or fragments thereof, wherein at least one valine (Val) of said variant, or fragments thereof is substituted with an amino acid selected from the group of amino acids consisting of Gly, Ala, Leu, and Ile, and independently thereof, variants, or fragments thereof, wherein at least one leucine (Leu) of said variant, or fragments thereof is substituted with an amino acid selected from the group of amino acids consisting of Gly, Ala, Val, and Ile, and independently thereof, variants, or fragments thereof, wherein at least one isoleucine (Ile) of said variants, or fragments thereof is substituted with an amino acid selected from the group of amino acids consisting of Gly, Ala, Val and Leu, and independently thereof, variants, or fragments thereof wherein at least one aspartic acids (Asp) of said variant, or fragments thereof is substituted with an amino acid selected from the group of amino acids consisting of Glu, Asn, and Gln, and independently thereof, variants, or fragments thereof, wherein at least one aspargine (Asn) of said variants, or fragments thereof is substituted with an amino acid selected from the group of amino acids consisting of Asp, Glu, and Gln, and independently thereof, variants, or fragments thereof, wherein at least one glutamine (Gln) of said variants, or fragments thereof is substituted with an amino acid selected from the group of amino acids consisting of Asp, Glu, and Asn, and wherein at least one phenylalanine (Phe) of said variants, or fragments thereof is substituted with an amino acid selected from the group of amino acids consisting of Tyr, Trp, His, Pro, and preferably selected from the group of amino acids consisting of Tyr and Trp, and independently thereof, variants, or fragments thereof, wherein at least one tyrosine (Tyr) of said variants, or fragments thereof is substituted with an amino acid selected from the group of amino acids consisting of Phe, Trp, His, Pro, preferably an amino acid selected from the group of amino acids consisting of Phe and Trp, and independently thereof, variants, or fragments thereof, wherein at least one arginine (Arg) of said fragment is substituted with an amino acid selected from the group of amino acids consisting of Lys and His, and independently thereof, variants, or fragments thereof, wherein at least one lysine (Lys) of said variants, or fragments

thereof is substituted with an amino acid selected from the group of amino acids consisting of Arg and His, and independently thereof, variants, or fragments thereof, and independently thereof, variants, or fragments thereof, and wherein at least one proline (Pro) of said variants, or fragments thereof is substituted with an amino acid selected from the group of amino acids consisting of Phe, Tyr, Trp, and His, and independently thereof, variants, or fragments thereof, wherein at least one cysteine (Cys) of said variants, or fragments thereof is substituted with an amino acid selected from the group of amino acids consisting of Asp, Glu, Lys, Arg, His, Asn, Gln, Ser, Thr, and Tyr.

It thus follows from the above that the same functional equivalent of a peptide fragment, or fragment of said functional equivalent may comprise more than one conservative amino acid substitution from more than one group of conservative amino acids as defined herein above. The term "conservative amino acid substitution" is used synonymously herein with the term "homologous amino acid substitution".

The groups of conservative amino acids are as the following:

P, A, G, S, T (neutral, weakly hydrophobic)

Q, N, E, D, B, Z (hydrophilic, acid amine)

H, K, R (hydrophilic, basic)

F, Y, W (hydrophobic, aromatic)

L, I, V, M (hydrophobic)

C (cross-link forming)

Conservative substitutions may be introduced in any position of a preferred predetermined peptide of the invention or fragment thereof. It may however also be desirable to introduce non-conservative substitutions, particularly, but not limited to, a non-conservative substitution in any one or more positions.

A non-conservative substitution leading to the formation of a functionally equivalent fragment of the peptide of the invention would for example differ substantially in polarity, for example a residue with a non-polar side chain (Ala, Leu, Pro, Trp, Val, Ile, Leu, Phe or Met) substituted for a residue with a polar side chain such as Gly, Ser, Thr, Cys, Tyr, Asn, or Gln or a charged amino acid such as Asp, Glu, Arg, or Lys, or substituting a charged or a polar residue for a non-polar one; and/or ii) differ

substantially in its effect on peptide backbone orientation such as substitution of or for Pro or Gly by another residue; and/or iii) differ substantially in electric charge, for example substitution of a negatively charged residue such as Glu or Asp for a positively charged residue such as Lys, His or Arg (and vice versa); and/or iv) differ substantially in steric bulk, for example substitution of a bulky residue such as His, Trp, Phe or Tyr for one having a minor side chain, e.g. Ala, Gly or Ser (and vice versa).

Substitution of amino acids may in one embodiment be made based upon their hydrophobicity and hydrophilicity values and the relative similarity of the amino acid side-chain substituents, including charge, size, and the like. Exemplary amino acid substitutions which take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

The addition or deletion of an amino acid may be an addition or deletion of from 2 to preferably 10 amino acids, such as from 2 to 8 amino acids, for example from 2 to 6 amino acids, such as from 2 to 4 amino acids. However, additions of more than 10 amino acids, such as additions from 2 to 10 amino acids, are also comprised within the present invention. In the multimeric forms additions/deletions may be made individually in each monomer of the multimer.

The invention also concerns non-peptide variants of the compounds disclosed herein. In particular, such variants should be understood to be compounds which bind to or in other ways interact with the Ig1, Ig2 or the Ig3 modules of NCAM and thereby stimulate Ig1, Ig2 or Ig3 signalling and/or modulate proliferation and/or induce differentiation and/or stimulate regeneration, neuronal plasticity and/or survival of cells presenting an NCAM receptor.

Functional equivalent

A functional equivalent may be obtained by substitution an amino acid in the sequence, which has a functional activity according to the invention. Functionally similar in the present content refers to dominant characteristics of the side chains such as hydrophobic, basic, neutral or acidic, or the presence or absence of steric

bulk. Accordingly, in one embodiment of the invention, the degree of identity between i) a given functional equivalent capable of effect and ii) a preferred predetermined fragment, is not a principal measure of the fragment as a variant or functional equivalent of a preferred predetermined peptide fragment according to the present invention.

Fragments sharing at least some homology with a preferred predetermined fragment of at least 3 amino acids, more preferably at least 5 amino acids, are to be considered as falling within the scope of the present invention when they are at least about 25 percent homologous with the preferred predetermined NCAM peptide, or fragment thereof, such as at least about 30 percent homologous, for example at least about 40 percent homologous, such as at least about 50 percent homologous, for example at least about 55 percent homologous, such as at least about 60 percent homologous, for example at least about 65 percent homologous, such as at least about 70 percent homologous, such as at least about 75 percent homologous, for example at least about 80 percent homologous, such as at least about 85 percent homologous.

Sequence identity can be measured using sequence analysis software (for example, the Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Centre, 1710 University Avenue, Madison, WI 53705), with the default parameters as specified therein.

Where nothing is specified it is to be understood that the C-terminal amino acid of a polypeptide of the invention exists as the free carboxylic acid, this may also be specified as "-OH". However, the C-terminal amino acid of a compound of the invention may be the amidated derivative, which is indicated as "-NH_2". Where nothing else is stated the N-terminal amino acid of a polypeptide comprise a free amino-group, this may also be specified as "H-".

Where nothing else is specified amino acid can be selected from any amino acid, whether naturally occurring or not, such as alpha amino acids, beta amino acids, and/or gamma amino acids. Accordingly, the group comprises but are not limited to: Ala, Val, Leu, Ile, Pro, Phe, Trp, Met, Gly, Ser, Thr, Cys, Tyr, Asn, Gln, Asp, Glu, Lys, Arg,His Aib, Nal, Sar, Orn, Lysine analogues DAP and DAPA, 4Hyp

Method for testing

According to the present invention, compounds capable of modulating the interaction between two individual NCAM molecules through the homophylic binding site composed by the Ig1, Ig2 and Ig3 modules of said NCAM molecules may be identified by testing their capability of

i) interacting with the Ig1 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig1 and Ig3 modules of NCAM, wherein said modules are from the two individual fragments of (b) interacting to each other, and/or

ii) interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig1 modules of NCAM, wherein said modules are from the two individual fragments of (b) interacting to each other, and/or

iii) interacting with the Ig2 module of NCAM, and thereby mimicking the interaction between Ig2 and Ig3 modules of NCAM, wherein said modules are from the two individual fragments of (b) interacting to each other, and/or

iv) interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig2 modules of NCAM, wherein said modules are from the two individual fragments of (b) interacting to each other, and/or

v) interacting with the Ig2 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig2 and Ig2 modules of NCAM, wherein said modules are from the two individual fragments of (b) interacting to each other

The method according to the invention comprises the steps of

a) providing a compound;

b) providing at least one individual fragment of an NCAM molecule, wherein said fragment comprises a sequence of consecutive amino acid residues corresponding to the sequence of the Ig1-2-3 module of NCAM comprising residues 1 to 289 of the sequence set forth in SEQ ID NO: 40 or fragments of said sequence;

c) contacting the compound of (a) with the individual NCAM fragment of (b), and

d) testing whether the compound is capable of

- interacting with the Ig1 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig1 and Ig3 modules of NCAM, wherein said modules are from the two individual fragments of (b) interacting to each other, and/or
- interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig1 modules of NCAM, wherein said modules are from the two individual fragments of (b) interacting to each other, and/or
- interacting with the Ig2 module of NCAM, and thereby mimicking the interaction between Ig2 and Ig3 modules of NCAM, wherein said modules are from the two individual fragments of (b) interacting to each other, and/or
- interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig2 modules of NCAM, wherein said modules are from the two individual fragments of (b) interacting to each other, and/or
- interacting with the Ig2 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig2 and Ig2 modules of NCAM, wherein said modules are from the two individual fragments of (b) interacting to each other.

d) providing a candidate compound capable of any of the above interaction(s).

In a preferred embodiment of the invention the individual fragment of NCAM is represented by the Ig1-2-3 module of NCAM comprising a consecutive sequence of at least 289 amino acids from the sequence of NCAM. In more preferred embodiment the sequence comprises aa 1 to 289 of NCAM, wherein NCAM is rat NCAM having the NCBI accession number NP_113709 identified as SEQ ID NO: 40 of the present application.

By the "Ig1-2-3 module of NCAM" in the present context is meant a contiguous amino acid sequence as described above consisting of the sequences of Ig1, Ig2, and Ig3, and linker sequences connecting said modules in the following order: N-terminus<Ig1-linker-Ig2-linker-Ig3 >C-terminus. The Ig1-2-3 module may be a

recombinant molecule consisting of the Ig1, Ig2 and Ig3 modules, or it may be a recombinant fusion protein containing the Ig1, Ig2 and Ig3 modules and a fusion partner, or it may be a fragment of a full-length NCAM molecule obtained by any known in the art method. According to the above invention the Ig1-2-3 module of the above method is in solution. In one embodiment the solution is an aquatic solution. In a a preferred embodiment the solution is phosphate buffered saline (PBS) solution or a TRIS-HCl buffer, pH 7.4.

The contacting of a compound with the individual NCAM fragment of step (c) preferably occurs in solution of the Ig1-2-3 module.

Testing whether the compound is capable of the above interactions may be done by any available in the art method currently used for the detection of protein interactions. For example the NMR spectroscopy may be selected as an appropriate method, or it may be done by using the Plasmon Resonance Analysis. According to the invention the NMR evaluation is preferred.

A compound capable of the above interaction(s) is identified by the above method is designated according to the invention as a candidate compound. The candidate compound may further be tested for its capability of modulating the interaction between at least two individual modules of NCAM, such as

i) the Ig1 module and the Ig3 module, and/or
ii) the Ig2 module and the Ig3 module, and/or
iii) the Ig2 module and the Ig2 module.

The latter testing may for example be done by using gel-filtration of the Ig1-2-3 module in solution in the presence of the selected candidate compound. Or it may be done by using gel-filtration of a mixture of the individual Ig1 and Ig3 modules, or a mixture of the individual Ig2 and Ig3 modules, or a mixture of two individual NCAM fragments, each of which is consisting of the contiguous sequence of the Ig1 and Ig3 modules, or the fragments consisting of Ig2 and Ig3 modules. Gel-filtration chromatography is one of the most commonly used laboratory techniques and a skilful artisan can easily perform such a testing.

According to the present invention, the candidate compound may be any molecule capable of modulating the interactions of the Ig1, Ig2 and Ig3 NCAM modules of above. Such a compound may, for example, be selected from the group comprising combinatorial libraries of peptides, lipids, carbohydrates or other organic molecules, or co-polymers of amino acids with other organic compounds. In a preferred embodiment, the candidate compound of the invention is a peptide.

The purpose of the above testing method is the identification and selection of interesting compounds (candidate compounds) capable of interacting with the Ig1-2-3 module of NCAM at the binding site of the invention and thereby modulating NCAM-dependent cell differentiation, adhesion and/or survival.

Crystal

According to the invention the identification of a candidate compound may comprise the use of a crystalline protein comprising either the individual Ig1, Ig2 and Ig3 modules or a combination of said modules.

A crystalline protein of the Ig1-2-3 module of NCAM consisting of the amino acid sequence corresponding to amino acid residues 1-289 of rat NCAM (SwissProt accession number NP_113709) (SEQ ID NO: 40) is made by the authors of the present invention to determine the structure of NCAM homophylic binding site and the computer generated 3D structure of the module is proposed herein for the in-silico screening compounds capable of binding to the identified homophylic binding site.

In a preferred embodiment the crystalline protein of the invention is a crystal of a polypeptide comprising the Ig1-2-3 module of NCAM comprising a homophilic binding site of NCAM. The crystal may comprise more than one polypeptide, for example two polypeptides. In a preferred embodiment the crystal comprises the Ig1, Ig2 and Ig3 modules of NCAM co-jointed in one fragment by interconnecting amino acid sequences, said one fragment termed herein "the Ig1-2-3 fragment".

Hence, it is preferred that the crystal diffracts X-rays for determination of atomic co-ordinates to a resolution of at least 4 Å, preferably at least 3 Å, more preferably at least 2.8 Å, even more preferably at least 2.5 Å, most preferably at least 2.0 Å.

In a very preferred embodiment of the invention the crystal comprises atoms arranged in a spatial relationship represented by the structure co-ordinates of table 2 shown on Figure 2, or by co-ordinates having a root mean square deviation there from of not more than 2.5 Å, preferably not more than 2.25 Å, more preferably not more than 2.0 Å, even more preferably not more than 1.75 Å, yet more preferably not more than 1.5 Å, for example not more than 1.25 Å, such as not more than 1.0 Å. Preferably, the co-ordinates has a root mean square deviation there from, of not more than 2.5 Å, preferably not more than 2.25 Å, more preferably not more than 2.0 Å, even more preferably not more than 1.75 Å, yet more preferably not more than 1.5 Å, for example not more than 1.25 Å, such as not more than 1.0 Å.

Preferably, the crystal comprises or more preferably consists of the structure as deposited to the PDB with id 1QZ1.

The crystal may comprise more than one polypeptide of the Ig1-2-3 fragment NCAM per asymmetric unit, in a preferred embodiment of the invention the crystal comprises polypeptides of the one Ig1-2-3 module of NCAM per asymmetric unit.

It is preferred that the crystal has unit cell dimensions of in the range of
a=50 to 52, preferably 50.5 to 51.0, more preferably around 51.5
b=107.5 to 109.5, preferably 108 to 109, more preferably around 108.5
c=146 to 151, preferably 148 to 150, more preferably around 149.0
α=85.5 to 95.5, preferably 88 to 92, more preferably around 90
β=85.5 to 95.5, preferably 88 to 92, more preferably around 90
γ=85.5 to 95.5, preferably 88 to 92, more preferably around 90.

Most preferably the crystal has the following characteristics:

Spacegroup: $I2_12_12_1$ with 1 molecule per asymmetric unit,
unit cell dimensions of a=51.5 b=108.5 c=149.0 Å alpha=90° beta=90° gamma=90°.

Preparing crystals

After several unsuccessful attempts, suitable conditions for preparing crystals of a polypeptide corresponding to the Ig 1-2-3 module of NCAM were identified.

It is therefore also an aspect of the present invention to provide a crystal comprising a polypeptide comprising at least 289 consecutive amino acid residues corresponding to amino acid residues 1-289 of rat NCAM (NCBI accession number NP_113709) (SEQ ID NO: 40), said consecutive amino acids correspond to the Ig1-2-3 fragment of rat NCAM using a method of preparing a crystal, wherein said method comprises the steps of

i) providing said polypeptide;

ii) growing crystals under conditions wherein said polypeptide is incubated in a buffer comprising in the range of 14 to 17% polyethylene glycol 4000 (PEG4k), in the range of 0.150 M to 0.5 M Li sulfate salt wherein said buffer has a pH in the range of 4.8 – 5.8;

iii) thereby preparing said crystals

In one embodiment of the invention, co-crystals of said polypeptide and a compound capable of interacting with said polypeptide are prepared. Said compound may have been identified by any of the methods outlined herein below. Hence, the compound may in one aspect of the invention be a modulator, such as a modulator of NCAM-homophilic interaction mediated by the Ig 1-2-3 module of NCAM.

The co-crystals are useful for designing optimised compounds, with enhanced binding properties. In particular, the co-crystals may be useful for designing better inhibitors of homophilic interaction mediated by the Ig 1-2-3 module of NCAM, or stabilizers of said interaction.

The buffer preferably comprises in the range of 5 to 25% polyethylene glycol, more preferably in the range of 10 to 20%, even more preferably in the range of 12 to 18%, yet more preferably in the range of 14 to 16 %, most preferably around 15% polyethylene glycol. Polyethylene glycol (PEG) may be any suitable PEG for example a PEG selected from the group consisting of PEG 4000, PEG 6000 and PEG 8000, preferably polyethylene glycol is PEG 4000.

The buffer preferably comprises in the range of 0.15 M to 0.5 M salt, more preferably in the range of 0.2 to 0.5 M, even more preferably in the range of 0.3 to 0.5 M, yet more preferably in the range of 0.4 to 0.5 M, most preferably around 0.45 M salt. The salt may be any useful salt, preferably the salt is Li sulfate (Li_2SO_4)

The buffer preferably has a pH in the range of 4.0 to 8.5, more preferably in the range of 4.5 to 7.5, even more preferably in the range of 5.0 to 6.5, yet more preferably in the range of 5.0 to 5.2. The buffer may be any useful buffer, preferably the Na-acetate buffer.

Incubation should be performed at a suitable temperature, preferably at a temperature in the range of 5 to 25°C, more preferably in the range of 10 to 25°C, even more preferably in the range of 15 to 25°C, even more preferably in the range of 17 to 21°C, yet more preferably around 18°C.

The crystals may be grown by any suitable method, for example by the hanging drop method.

Determination of structure

The structure of crystals may be determined by any method known to the person skilled in the art, for example using X-ray diffraction. Once a structure has been identified, said structure may be refined using suitable software.

In one embodiment of the invention a molecular replacement technique may be used. Such techniques involves that the structure is determined by obtaining x-ray diffraction data for crystals of the polypeptide or complex for which one wishes to determine the three dimensional structure. Then, one determines the three-dimensional structure of that polypeptide or complex by analysing the x-ray diffraction data using molecular replacement techniques with reference to known structural co-ordinates of a structurally similar protein. In the case of polypeptide comprising the Ig1-2 modules of NCAM, structural co-ordinates of said modules may be used. As described in U.S. Pat. No. 5,353,236, for instance, molecular replacement uses a molecule having a known structure as a starting point to model the structure of an unknown crystalline sample. This technique is based on the principle that two molecules, which have similar structures, orientations and

positions in the unit cell, diffract similarly. Molecular replacement involves positioning the known structure in the unit cell in the same location and orientation as the unknown structure. Once positioned, the atoms of the known structure in the unit cell are used to calculate the structure factors that would result from a hypothetical diffraction experiment. This involves rotating the known structure in the six dimensions (three angular and three spatial dimensions) until alignment of the known structure with the experimental data is achieved. This approximate structure can be fine-tuned to yield a more accurate and often higher resolution structure using various refinement techniques. For instance, the resultant model for the structure defined by the experimental data may be subjected to rigid body refinement in which the model is subjected to limited additional rotation in the six dimensions yielding positioning shifts of under about 5%. The refined model may then be further refined using other known refinement methods.

Another method for determining the three-dimensional structure of a polypeptide corresponding to the Ig 1-2-3 module of NCAM, or a complex of said polypeptide with an interacting compound, is homology modelling techniques. Homology modelling involves constructing a model of an unknown structure using structural co-ordinates of one or more related proteins, protein domains and/or subdomains. Homology modelling may be conducted by fitting common or homologous portions of the protein or peptide whose three dimensional structure is to be solved to the three dimensional structure of homologous structural elements. Homology modelling can include rebuilding part or all of a three dimensional structure with replacement of amino acids (or other components) by those of the related structure to be solved.

An example of structure determination is outlined in example 2.

Structural coordinates of a crystalline polypeptide of this invention may be stored in a machine-readable form on a machine-readable storage medium, e.g. a computer hard drive, diskette, DAT tape, CD-ROM etc., for display as a three-dimensional shape or for other uses involving computer-assisted manipulation of, or computation based on, the structural coordinates or the three-dimensional structures they define. For example, data defining the three dimensional structure of a polypeptide corresponding to the Ig 1-2-3 module of NCAM, may be stored in a machine-readable storage medium, and may be displayed as a graphical three-dimensional

representation of the protein structure, typically using a computer capable of reading the data from said storage medium and programmed with instructions for creating the representation from such data. This invention thus encompasses a machine, such as a computer, having a memory that contains data representing the structural coordinates of a crystalline composition of this invention, e.g. the coordinates set forth in table 2 (Figure 2), together with additional optional data and instructions for manipulating such data. Such data may be used for a variety of purposes, such as the elucidation of other related structures and drug discovery.

A first set of such machine readable data may be combined with a second set of machine-readable data using a machine programmed with instructions for using the first data set and the second data set to determine at least a portion of the coordinates corresponding to the second set of machine-readable data. For instance, the first set of data may comprise a Fourier transform of at least a portion of the coordinates for the complex set forth in table 2 (Figure 2), while the second data set may comprise X-ray diffraction data of a molecule or molecular complex.

More specifically, one of the objects of this invention is to provide three-dimensional structural information of co-complexes comprising the homophilic binding site of the Ig 1-2-3 module of NCAM. To that end, we provide for the use of the structural co-ordinates of a crystalline composition of this invention, or portions thereof, to solve, e.g. by molecular replacement or by homology modelling techniques, the three dimensional structure of a crystalline form of another similar cell adhesion molecule (CAM), for example another CAM comprising the Ig modules capable of homophilic interaction or a polypeptide:interacting compound complex.

For example, one may use molecular replacement to exploit a set of coordinates such as set forth in table 2 (Figure 2) to determine the structure of a crystalline co-complex of a polypeptide corresponding to the Ig 1-2-3 module of NCAM comprising a homophilic binding site and an interacting compound.

Uses of the structures

A 3D representation of the polypeptides described in the present invention may be useful for several purposes, for example for determining the structure of similar proteins or polypeptides (see also herein above) or for designing compounds capable of interacting with said polypeptides.

For example, the three dimensional structure defined by the machine readable data for the polypeptide of the Ig1-2-3 module of NCAM may be computationally evaluated for its ability to associate with various chemical entities or test compounds. The term "chemical entity", as used herein, refers to chemical compounds, complexes of at least two chemical compounds, and fragments of such compounds or complexes.

For instance, a first set of machine-readable data defining the 3-D structure of polypeptide corresponding to the Ig1-2-3 module of NCAM or complex thereof, is combined with a second set of machine-readable data defining the structure of a chemical entity or test compound of interest using a machine programmed with instructions for evaluating the ability of the chemical entity or compound to associate with the Ig1-2-3 module of NCAM or complex thereof and/or the location and/or orientation of such association. Such methods provide insight into the location, orientation and energies of association of protein surfaces with such chemical entities.

The three dimensional structure defined by the data may be displayed in a graphical format permitting visual inspection of the structure, as well as visual inspection of the association of the polypeptide component(s) with an interacting compound. Alternatively, more quantitative or computational methods may be used. For example, one method of this invention for evaluating the ability of a chemical entity to associate with any of the molecules or molecular complexes set forth herein comprises the steps of: (a) employing computational means to perform a fitting operation between the chemical entity and a binding site or other surface feature of the molecule or molecular complex; and (b) analysing the results of said fitting operation to quantify the association between the chemical entity and the binding site.

This invention further provides for the use of the structural coordinates of a crystalline composition of this invention, or portions thereof, to identify reactive amino acids, such as cysteine residues, within the three-dimensional structure, preferably within or adjacent to a binding site; to generate and visualise a molecular surface, such as a water-accessible surface or a surface comprising the space-filling van der Waals surface of all atoms; to calculate and visualise the size and shape of surface features of the protein or complex, e.g., substrate binding sites; to locate potential H-bond donors and acceptors within the three-dimensional structure, preferably within or adjacent to a ligand binding site; to calculate regions of hydrophobicity and hydrophilicity within the three-dimensional structure, preferably within or adjacent to a ligand binding site; and to calculate and visualize regions on or adjacent to the protein surface of favourable interaction energies with respect to selected functional groups of interest (e.g. amino, hydroxyl, carboxyl, methylene, alkyl, alkenyl, aromatic carbon, aromatic rings, heteroaromatic rings, etc.). One may use the foregoing approaches for characterising the polypeptide corresponding to the Ig1-2-3 module of NCAM and its interactions with moieties of potential interacting compounds to design or select compounds capable of specific covalent attachment to reactive amino acids (e.g., cysteine) and to design or select compounds of complementary characteristics (e.g., size, shape, charge, hydrophobicity/hydrophilicity, ability to participate in hydrogen bonding, etc.) to surface features of the protein, a set of which may be preselected. Using the structural coordinates, one may also predict or calculate the orientation, binding constant or relative affinity of a given ligand to the protein in the complexed state, and use that information to design or select compounds of improved affinity.

In such cases, the structural coordinates of the polypeptide of the Ig1-2-3 module of NCAM, or portion or complex thereof, are entered in machine readable form into a machine programmed with instructions for carrying out the desired operation and containing any necessary additional data, e.g. data defining structural and/or functional characteristics of a potential interacting compound or moiety thereof, defining molecular characteristics of the various amino acids, etc.

One method of this invention provides for selecting from a database of chemical structures a compound capable of binding to the Ig1-2-3 module of NCAM. The method starts with structural co-ordinates of a crystalline composition of the

invention, e.g., co-ordinates defining the three dimensional structure of the Ig 1-2-3 module of NCAM or a portion thereof or a complex thereof. Points associated with that three-dimensional structure are characterised with respect to the favourable ability of interactions with one or more functional groups. A database of chemical structures is then searched for candidate compounds containing one or more functional groups disposed for favourable interaction with the protein based on the prior characterisation. Compounds having structures which best fit the points of favourable interaction with the three dimensional structure are thus identified.

It is often preferred, although not required, that such searching be conducted with the aid of a computer. In that case a first set of machine-readable data defining the 3D structure of a polypeptide corresponding to the Ig1-2-3 module of NCAM, or a portion or polypeptide/interacting compound complex thereof, is combined with a second set of machine readable data defining one or more moieties or functional groups of interest, using a machine programmed with instructions for identifying preferred locations for favourable interaction between the functional group(s) and atoms of the polypeptide. A third set of data, i.e. data defining the location(s) of favourable interaction between polypeptide and functional group(s) is so generated. That third set of data is then combined with a fourth set of data defining the 3D structures of one or more chemical entities using a machine programmed with instructions for identifying chemical entities containing functional groups so disposed as to best fit the locations of their respective favourable interaction with the polypeptide.

Compounds having the structures selected or designed by any of the foregoing means may be tested for their ability to bind to the Ig 1-2-3 module of NCAM.

In one preferred embodiment of the invention, the compound is preferably a modulator of NCAM homophilic interaction mediated by the Ig 1-2-3 fragment. For example, a compound capable of interacting with the Ig1-2-3 homophilic binding site may be a good inhibitor of NCAM homophilic binding and NCAM function that requires this binding. Hence, compounds having the structures selected or designed by any of the foregoing means may be tested for their ability to modulate NCAM activity, such as mediation of cell differentiation and/or survival of NCAM presenting cells.

As practitioners in this art will appreciate, various computational analyses may be used to determine the degree of similarity between the three dimensional structure of a given polypeptide (or a portion or complex thereof) and a polypeptide corresponding to the Ig1-2-3 module of NCAM or complex thereof such as are described herein. Such analyses may be carried out with commercially available software applications, such as the Molecular Similarity application of QUANTA (Molecular Simulations Inc., Waltham, Mass.) version 3.3, and as described in the accompanying User's Guide, Volume 3 pgs. 134-135.

The Molecular Similarity application permits comparisons between different structures, different conformations of the same structure, and different parts of the same structure. The procedure used in Molecular Similarity to compare structures is divided into four steps: (1) load the structures to be compared; (2) define the atom equivalences in these structures; (3) perform a fitting operation; and (4) analyse the results.

Each structure is identified by a name. One structure is identified as the target (i.e., the fixed structure); all remaining structures are working structures (i.e., moving structures). Since atom equivalency within QUANTA is defined by user input, for the purpose of this invention we define equivalent atoms as protein backbone atoms (N, Cα, C and O) for all conserved residues between the two structures being compared and consider only rigid fitting operations.

When a rigid fitting method is used, the working structure is translated and rotated to obtain an optimum fit with the target structure. The fitting operation uses a least squares fitting algorithm that computes the optimum translation and rotation to be applied to the moving structure, such that the root mean square difference of the fit over the specified pairs of equivalent atom is an absolute minimum. This number, given in angstroms, is reported by QUANTA.

For the purpose of this invention, any set of structural co-ordinates of a polypeptide corresponding to Ig 1-2-3 module of NCAM or molecular complex thereof that has a root mean square deviation of conserved residue backbone atoms (N, Cα, C, O) of less than 1.5 Å when superimposed--using backbone atoms--on the relevant

structural co-ordinates of a protein or complex of this invention, e.g. the co-ordinates listed in table 2 (Figure 2), are considered identical. More preferably, the root mean square deviation is less than 1.0 Å. Most preferably, the root mean square deviation is less than 0.5 Å.

The term "root mean square deviation" means the square root of the arithmetic mean of the squares of the deviations from the mean. It is a way to express the deviation or variation from a trend or object. For purposes of this invention, the "root mean square deviation" defines the variation in the backbone of a protein from the backbone of a protein of this invention, such as a homophilic binding site of the Ig 1-2-3 module of NCAM as defined by the structural co-ordinates of table 2 (Figure 2) and described herein.

The term "least squares" refers to a method based on the principle that the best estimate of a value is that in which the sum of the squares of the deviations of observed values is a minimum.

In order to use the structural co-ordinates generated for a crystalline substance of this invention, e.g. the structural co-ordinates set forth in table 2 (Figure 2), it is often necessary or desirable to display them as, or convert them to, a three-dimensional shape, or to otherwise manipulate them. This is typically accomplished by the use of commercially available software such as a program, which is capable of generating three-dimensional graphical representations of molecules or portions thereof from a set of structural co-ordinates.

By way of illustration, a non-exclusive list of computer programs for viewing or otherwise manipulating protein structures include the following:
Midas (Univ. of California, San Francisco),
MidasPlus (Univ. of Cal., San Francisco)
MOIL (Univeristy of Illinois)
Yummie (Yale University)
Sybyl (Tripos, Inc.)
Insight/Discover (Biosym Technologies)
MacroModel (Columbia University)
Quanta (Molecular Simulations, Inc.)

Cerius (Molecular Simulations, Inc.)
Alchemy (Tripos, Inc.)
LabVision (Tripos, Inc.)
Rasmol (Glaxo Research and Development)
Ribbon (University of Alabama)
NAOMI (Oxford University)
Explorer Eyechem (Silicon Graphics, Inc.)
Univision (Cray Research)
Molscript (Uppsala University)
Chem-3D (Cambridge Scientific)
Chain (Baylor College of Medicine)
O (Uppsala University)
GRASP (Columbia University)
X-Plor (Molecular Simulations, Inc.; Yale Univ.)
Spartan (Wavefunction, Inc.)
Catalyst (Molecular Simulations, Inc.)
Molcadd (Tripos, Inc.)
VMD (Univ.of Illinois/Beckman Institute)
Sculpt (Interactive Simulations, Inc.)
Procheck (Brookhaven Nat'l Laboratory)
DGEOM (QCPE)
RE_VIEW (Brunel University)
Modeller (Birbeck Col., Univ. of London)
Xmol (Minnesota Supercomputing Center)
Protein Expert (Cambridge Scientific)
HyperChem (Hypercube)
MD Display (University of Washington)
PKB (Nat'l Center for Biotech. Info., NIH)
ChemX (Chemical Design, Ltd.)
Cameleon (Oxford Molecular, Inc.)
Iditis (Oxford Molecular, Inc.)

For storage, transfer and use with such programs of structural coordinates for a crystalline substance of this invention, a machine-readable storage medium is provided comprising a data storage material encoded with machine readable data

which, when using a machine programmed with instructions for using said data, e.g. a computer loaded with one or more programs of the sort identified above, is capable of displaying a graphical three-dimensional representation of any of the molecules or molecular complexes described herein. Machine-readable storage media comprising a data storage material include conventional computer hard drives, floppy disks, DAT tape, CD-ROM, and other magnetic, magneto-optical, optical, floptical and other media which may be adapted for use with a computer.

Even more preferred is a machine-readable data storage medium that is capable of displaying a graphical three-dimensional representation of a molecule or molecular complex that is defined by the structural co-ordinates of the Ig1-2-3 module of NCAM, such as the co-ordinates set forth in table 2 (Figure 2) +/- a root mean square deviation from the conserved backbone atoms of the amino acids thereof of not more than 1.5 Å. An illustrative embodiment of this aspect of the invention is a conventional 3.5" diskette, DAT tape or hard drive encoded with a data set, preferably in PDB format, comprising the co-ordinates of table 2 (Figure 2). FIG. 3 illustrates a print-out of a graphical three-dimensional representation of such a polypeptide.

In another embodiment, the machine-readable data storage medium comprises a data storage material encoded with a first set of machine readable data which comprises the Fourier transform of the structural co-ordinates set forth in table 2 (Figure 2) (or again, a derivative thereof), and which, when using a machine programmed with instructions for using said data, can be combined with a second set of machine readable data comprising the X-ray diffraction pattern of a molecule or molecular complex to determine at least a portion of the structural co-ordinates corresponding to the second set of machine readable data.

Such a system may for example include a computer comprising a central processing unit ("CPU"), a working memory which may be, e.g., RAM (random-access memory) or "core" memory, mass storage memory (such as one or more disk drives or CD-ROM drives), one or more cathode-ray tube ("CRT") display terminals, one or more keyboards, one or more input lines (IP), and one or more output lines (OP), all of which are interconnected by a conventional bidirectional system bus.

Input hardware, coupled to the computer by input lines, may be implemented in a variety of ways. Machine-readable data of this invention may be inputted via the use of a modem or modems connected by a telephone line or dedicated data line. Alternatively or additionally, the input hardware may comprise CD-ROM drives or disk drives. In conjunction with the CRT display terminal, a keyboard may also be used as an input device.

Output hardware, coupled to the computer by output lines, may similarly be implemented by conventional devices. By way of example, output hardware may include a CRT display terminal for displaying a graphical representation of a protein of this invention (or portion thereof) using a program such as QUANTA as described herein. Output hardware might also include a printer, so that hard copy output may be produced, or a disk drive, to store system output for later use.

In operation, the CPU coordinates the use of the various input and output devices, co-ordinates data accesses from mass storage and accesses to and from working memory, and determines the sequence of data processing steps. A number of programs may be used to process the machine-readable data of this invention. Examples of such programs are discussed herein above. Algorithms suitable for this purpose are also implemented in programs such as Cast-3D (Chemical Abstracts Service), 3DB Unity (Tripos, Inc.), Quest-3D (Cambridge Crystallographic Data Center), and MACCS/ISIS-3D (Molecular Design Limited). These geometric searches can be augmented by steric searching, in which the size and shape requirements of the binding site are used to weed out hits that have prohibitive dimensions. Programs that may be used to synchronize the geometric and steric requirements in a search applied to the FRB of FRAP include CAVEAT (P. Bartlett, University of California, Berkeley), HOOK (MSI), ALADDIN (Daylight Software) and DOCK (http://www.cmpharm.ucsf.edu/kuntz-/kuntz.html and references cited therein). All of these searching protocols may be used in conjunction with existing corporate databases, the Cambridge Structural Database, or available chemical databases from chemical suppliers.

In one embodiment of the invention the methods involve identifying a number of compounds potentially capable of interacting with the Ig 1-2-3 module of NCAM or a fragment thereof, for example the methods may involve identification of a sub-library

of compounds potentially interacting with the Ig 1-2-3 module of NCAM or fragments thereof. This may be accomplished using any conventional method. For example, all the possible members of a combinatorial library may first be enumerated, according to the available reagents and the established synthetic chemistries. Individual members may then separately be docked into a binding site of a polypeptide of MASP-2. Finally, an optimal sub-library may be selected for synthesis, based on the ranking of their docking scores and/or diversity measures. Software for fast library enumeration has been developed, including for example CombiLibMaker in Sybyl, Analog Builder in Cerius2, and the QuaSAR-CombiGen module available in MOE (MOE Software, Chemical Computing Group, 1010 Sherbrooke Street W., Suite 910, Montreal, Canada H3A 2R7). Most of these programs can easily generate all of the 2D or 3D structures for a combinatorial library containing millions of compounds, using either fragment-based or reaction-based schemes. Other tools within these software packages are also available for decreasing the size of a virtual library prior to docking. For example, a library enumerated through CombiLibMaker can subsequently be analysed with diverse solutions (available in Sybyl) to provide a sub-library that adequately samples chemical space. QuaSAR-CombiDesign is another combinatorial library design tool available in MOE that provides a non-enumerative method for combinatorial library generation, and can, e.g. test against rule of five filters using statistical sampling techniques during library creation, creating smaller sub-libraries with user-defined property ranges. In principle, the docking step that follows library creation can be conducted using any of the available docking programs like DOCK or FlexX ©, while the diversity selection for example may be performed using software available from Daylight, Tripos (diverse solutions), or BCI or by high throughput docking as for example described by Diller and Merz.

In another example a 'divide-and-conquer' approach may be used. With this strategy, all of the product structures in a combinatorial library are viewed as having variable substituents attached through one or multiple sites on a common template. The template is first docked into the binding site and only the top-scoring poses are saved for the further consideration. Individual substituents are then independently attached onto each pose of the template, to assess which substituents can fit well into the binding site. Only those combinations of top-scoring substituents are further considered and scored to identify the whole product structures that can dock really

well into the binding site. This may be done with the aid of suitable software for example PRO SELECT, CombiBUILD, CombiDOCK, DREAM ++ and FlexX ©.

In one embodiment the methods of invention comprise application of pharmacophores obtained using active site maps. Herein the term "active site" is meant to describe a site responsible of interaction with a compound and not a catalytically active site. The method may for example be a computational approach comprising the generation of multiple, promising, structurally diverse test-compounds. The search for multiple structural series may be accomplished by coupling protein structural information with combinatorial library design using any suitable method. For example the "design in receptor" method (Murrary et al., 1999) or the method outlined herein below may be used. Methods to account for multiple protein conformations for example as described by Mason et al., 2000 may also be used, including the creation of a dynamic pharmacophore model (as for example described by Carlson et al., 2000) from molecular dynamics simulations. Also experimental and computational needle screening approaches for mapping active sites with molecular fragments may be used for example as described in Boehm et al., 2000. Any suitable software tools for mapping site points (e.g. GRID and SITEPOINT) may be used with the invention. Also MCSS techniques for generating site maps may be used.

Suitable methods may for example comprise generation of active site maps from protein structures. Then all possible 2-, 3- and 4-point pharmacophores can be enumerated from the site map and encoded as a bit string (signature) these pharmacophores define a space to be probed by compounds that are selected using the informative library design tool. The metric used to evaluate the success of the approach is the number of active scaffolds selected in the library design, with the number of active compounds as a secondary measure. Any suitable algorithm for site map generation may be used, for example algorithms generating between 10 and 80 feature positions for each active site. An example of such a method is described for example by Eksterowicz et al J Mol Graph Model. 2002 Jun;20(6):469-77.

Information of the various binding sites of the Ig1-2-3 module along with the crystal structure of the invention provide a tool for the examination of the biological

significance of the observed Ig1-to-Ig2, Ig1-to-Ig3, and Ig2-to-Ig3 contacts, and for the screening for compounds capable of mimicking the binding of the Ig1-to-Ig2, Ig1-to-Ig3, Ig3-to-Ig1, Ig2-to-Ig3 and Ig2-to-Ig2 modules of NCAM.

The structure of the Ig1-2-3 module in solution

Alternatively, the 3D structure of soluble Ig1-2-3 module may be determined and used for in-silico screening of the compounds for evaluation their potential to interact with the binding site comprised by the module. The NMR spectroscopy may ultimately be used for solving the structure of the proteins in solution.

Screening the compounds

The identification of a new compound capable of modulating adhesion, differentiation and/or survival of NCAM presenting cells may in one aspect be performed by screening a computer model template, such as for example the three-dimensional structure of the Ig1-2-3 module of NCAM as crystalline or soluble protein. Accordingly, the invention also relates to providing a screening method for selecting a compound capable of modulating cell differentiation and/or survival of NCAM presenting cells, comprising the steps of

i) providing a polypeptide comprising the Ig1-2-3 module of NCAM;
ii) preparing a crystalline protein comprising the polypeptide of (i);
iii) generating a structural model of the Ig1-2-3 module of NCAM as a crystalline protein of (ii);
iv) designing a compound into the structure of said generated model of step i);
v) selecting a compound capable of interacting with the homophylic binding site according to the structural model of (iii);
vi) testing the compound of step (vi) in an in vitro or in vivo assay whether the compound is capable of modulating neural cell differentiation, adhesion and/or survival.

The above screening method may in some embodiments comprise using a computer generated model of the Ig1-2-3 module of NCAM, or fragments of said module, such as Ig1, Ig2, Ig3, or Ig1-2, or Ig2-3 modules in a solution. Such a model may be generated on the basis of the data obtained, for example, from Nuclear Magnetic Resonance spectroscopy of samples of the above modules. However,

preferably, a computer generated model is a structural model of a crystal of the above modules.

The invention provides a computer generated structural structure model of the Ig1-2-3 module for the screening a compound capable of modulating NCAM homophylic binding-dependent cell adhesion, survival and differentiation.

Designing interacting compounds

Designing interacting compounds

Generating a site map

Feature points complementary to the active site are computed using an internally developed software tool. For example, a hydrogen bond donor feature is mapped in the proximity of a hydrogen bond acceptor in the protein active site. The collection of 3D coordinates and labels (acceptors, donors, negatives, positives, hydrophobes and aromatics) is called a site map. Technically, the site map is the union of three separately computed maps, ESMap which contains the electrostatic feature points (P, N, and H) HBMap with hydrogen-bonding feature points (D and A) and AroMap containing aromatic feature points (Ar).

The electrostatic feature map, ESMap, is computed by first using the sphere placement algorithm employed in the program PASS (Brady et al., 2000). It generates an evenly-distributed setof points (ProbeMap) in regions of buried volume along the protein surface. A subset of points in the ProbeMap comprises the P, N, and H feature points depending upon the local electrostatic character of the protein. The CVFF molecular mechanics force field is used to compute the electrostatic potential, φi , at each point *i* of ProbeMap, along with the mean potential φ and mean magnitude | φ | averaged over all points in ProbeMap. The value of φi determines whether or not point *i* is included as a P, N, or H feature point, according to the following definitions

$i>\varphi+1.5*\sigma(\varphi)$, i=N feature point

$i>\varphi-1.5*\sigma(\varphi)$, i=P feature point

$|\varphi|-1.0*\sigma(|\varphi|)<|\varphi i|<(|\varphi|)+1.0*\sigma(|\varphi|)$, i=H feature point

Here σ (X) denotes the standard deviation about the mean of quantity X. This normalizes the point assignments relative to the overall electrostatic environment of the active site. This presents non charge-neutral protein structures (which may result from counter ions not being resolved or present in the crystal structures) from skewing feature point assignments unreasonably.

The hydrogen-bonding feature map, HBMap, is determined by projecting complementary points outward from known hydrogen-bonding atoms of the protein. The resulting superset of points is filtered on the basis of steric clash, insufficient burial and minimal proximity of alike feature points. Ideal hydrogen-bonding points are positioned on the basis of the mean angle and distance as observed in the PDB (see for example table 2 shown on Figure 1). Points that clash with the protein are removed. However, for robustness, small positional perturbations are applied to retain potentially important hydrogen-bonding positions. Bifurcated hydrogen-bonding joints are computed heuristically by investigating full rings of points equally bifurcated between protein atoms that are considered moderate or strong hydrogen bond participants.

Points on such rings are retained as bifurcated HB points if they do not violate steric clash, burial and mutual proximity conditions. To build the final HBMap, the surviving sets of ideal and bifurcated HB points are combined and subjected to filtration on the basis of mutual proximity.

The AroMap set of aromatic feature points is computed by repeatedly docking a benzene ring into the protein active site and retaining the centroids of the top-scoring configurations. The protein is represented using a polar-hydrogen CVFF force field. The docking is performed using internal code in local optimization mode. One hundred separate local docking trials with different starting positions are performed. Any of the docked configurations whose score lies within an energy window of 5 kcal/mol of the minimum-energy configuration is included in AroMap. Again points are subjected to filtration on the basis of burial and mutual proximity.

Converting pharmacophores into a signature

Pharmacophores are generated on the basis of feature points in the active site by exhaustive enumeration of all 2-,3-, and 4-point subsets of the feature points. For all pairs of feature points their distance in 3D-space is precomputed. In order to arrive at a discrete representation of a pharmacophore, the distances are binned, applying

a user-defined binning scheme. Chirality is denoted by encoding the handedness of 4-point pharmacophores. Each pharmacophore is mapped onto a unique address, such that any possible combination of up to four features and distances is represented. The address is taken for a binary representation of the pharmacophores, called a signature. The length of the signature is the highest possible address for an encoding of a 4-point pharmacophore. All bits in the signature are initially set to 0. In order to represent a pharmacophore the bit at the respective address in the signature is turned on (set to 1). For the representation of the active site all pharmacophores are exhaustively enumerated and the respective bits are turned on.

Union of signatures for multiple structures

Multiple signatures may be combined. The binary union of multiple signatures yields a single bit string representing all pharmacophores present in any structure. Any consensus threshold *c* can be used to define the consensus representation of multiple active sites. That is, a pharmacophore is present in at least *c* of active site conformations. Note that this way of handling multiple active site snapshots is quite expedient.

Molecular signatures

Test compounds are encoded as follows. First, conformers are generated for each compound using an internal tool that generates a fairly complete conformational model of the molecule. Features are assigned using a substructure-based set of rules. Pharmacophores are enumerated from these three-dimensional feature positions following the same protocol as for the active site, thus ensuring compatibility of the binary encodings. However, multiple conformers need to be represented simultaneously here. This is done by wrapping the exhaustive enumeration of pharmacophores for a single conformer into an extra loop over all the conformers of a compound. That is, any pharmacophore on any conformer of a compound is represented by turning the respective bit in the signature on.

Molecular signature masking

With the binary representation of the active site and the binary representation of the molecules being defined analogously, the meaning of a bit at a certain address is the same (the same pharmacophore, within the tolerances of the distance binning).

Therefore, representing a design space amounts to masking all molecule signatures by the active site signature. Masking a signature means taking the logical *and* of the bits of the site signature and the molecule signature. For a given molecule, bits representing pharmacophores not present in the active site are turned off, whereas the bits of the pharmacophores in the active site can be either on or off, depending on their presence or absence in the molecules. This way only the pharmacophore space defined by the active site is taken into account.

Informative library design

Informative library design is a molecule selection strategy that optimises information return for a given virtual library. The goal is to detect a set of features (pharmacophores) that determine activity against a particular test compound. Informative design aims at selecting a set of compounds such that the resulting subset will interrogate the test compound in different, but overlapping ways. Molecules are selected for synthesis and screening such that each pharmacophore in the design space has a unique pattern of occurrence in the molecules of the set. This unique 'code' enables the identification and retention of the important pharmacophores when the set of compounds is assayed, regardless of the actual experimental outcome. This is in contrast to diversity methods that seek to produce a unique pattern of pharmacophore occurrences in each molecule.

Given a design space, the algorithm seeks to optimize decoding as many pharmacophores as possible, with the smoothest distribution across the size of pharmacophore classes. A pharmacophore class refers to the subset of pharmacophores that all have the same code or pattern. Note that the optimum solution is a set of compounds that enables decoding each individual pharmacophore. However, this may not be possible due either to the source pool, bit correlation or to limited size of selection. The cost function for an unconstrained optimisation in terms of molecule selection is the entropy of the class distribution. The entropy is given by

$$H - - \sum_{i=1}^{c} \frac{|Ci|}{f} \ln \frac{|Ci|}{f}$$

where *H* is the entropy of the feature classes, *C* the number of distinct classes, *f* the number of features in the design space and /c | is the size of class i. During the course of the optimisation, molecules are selected, such as to maximize H.

Testing

A compound selected by in-silico screening of the above is to be further tested whether it has a capability to modulate cellular adhesion, differentiation and/or survival of cells presenting NCAM.

Biological assays for the testing a capability of compounds to modulate adhesion, differentiation and/or survival of NCAM presenting cells are well known in the art. In particular, the assays described in WO03020749, WO0247719, WO03016351 and in the present application may be used.

The testing of the identified compound of (vi) may in some embodiments be additionally or alternatively performed by using the testing method described above.

Compound

A compound of the present invention is preferably selected by any of the above screening methods. According to the invention a selected compound is the candidate compound.

By the term "candidate compound" is meant a compound capable of

i) interacting with the Ig1 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig1 and Ig3 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

ii) interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig1 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

iii) interacting with the Ig2 module of NCAM, and thereby mimicking the interaction between Ig2 and Ig3 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

iv) interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig2 modules of NCAM, wherein said modules are from two individual NCAM molecules, and/or

v) interacting with the Ig2 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig2 and Ig2 modules of NCAM, wherein said modules are from two individual NCAM molecules.

A candidate compound, which is capable of at least one of the above interactions, is according to the invention also capable of modulating cell differentiation, adhesion and survival mediated by NCAM homophylic binding though the homophylic binding site described herein.

On one preferred embodiment a candidate compound is capable of interacting with the Ig1 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig1 and Ig3 modules of NCAM, wherein said modules are from two individual NCAM molecules, and thereby modulating cell differentiation, adhesion and survival mediated by NCAM homophylic binding through the homophylic binding site of the invention.

In another preferred embodiment a candidate compound is capable of interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig1 modules of NCAM, wherein said modules are from two individual NCAM molecules, and thereby modulating cell differentiation, adhesion and survival mediated by NCAM homophylic binding through the homophylic binding site of the invention.

A candidate compound, which is capable of interacting with the Ig2 module of NCAM, and thereby mimicking the interaction between Ig2 and Ig3 modules of NCAM, wherein said modules are from two individual NCAM molecules, and thereby modulating cell differentiation, adhesion and survival mediated by NCAM homophylic binding through the homophylic binding site of the invention is another preferred embodiment of the invention.

In still another preferred embodiment a candidate compound is capable of interacting with the Ig3 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig3 and Ig2 modules of NCAM, wherein said modules are from two individual NCAM molecules, and thereby modulating cell differentiation, adhesion and survival mediated by NCAM homophylic binding through the homophylic binding site of the invention.

In yet another preferred embodiment a candidate compound is capable of interacting with the Ig2 module of NCAM, and thereby mimicking and/or modulating the interaction between the Ig2 and Ig2 modules of NCAM, wherein said modules are from two individual NCAM molecules, and thereby modulating cell differentiation, adhesion and survival mediated by NCAM homophylic binding through the homophylic binding site of the invention.

Other preferred embodiments of the invention concern the

i) candidate compounds, which are capable of simultaneous interaction with amino acid residues of the Ig1 and Ig3 modules of the same NCAM molecule and thereby mimicking and/or modulating the interaction between the Ig1-to-Ig3 and the Ig3-to-Ig1 modules of two individual NCAM molecules interacting through said modules,

ii) candidate compounds, which are capable of simultaneous interaction with amino acid residues of the Ig1 and Ig2 modules of the same NCAM molecule and thereby mimicking and/or modulating the interaction the Ig1-to-Ig3 and the Ig2-to-Ig3 modules of two individual NCAM molecules interacting through said modules,

iii) candidate compounds, which are capable of simultaneous interaction with amino acid residues of the Ig1 and Ig3 modules of the same NCAM molecule and thereby mimicking and/or modulating the interaction between the Ig1-to-Ig3 and the Ig3-to-Ig2 modules of two individual NCAM molecules interacting through said modules,

iv) candidate compounds, which are capable of simultaneous interaction with amino acid residues of the Ig1 and Ig2 modules of the same NCAM molecule and thereby mimicking and/or modulating the interaction between the Ig1-to-Ig3 and the Ig2-to-Ig2 modules of two individual NCAM molecules interacting through said modules,

v) candidate compounds, which are capable of simultaneous interaction with amino acid residues of the Ig2 and Ig3 modules of the same NCAM molecule and thereby mimicking and/or modulating the interaction between the Ig2-to-Ig3 and the Ig3-to-Ig1 modules of two individual NCAM molecules interacting through said modules,

vi) candidate compounds, which are capable of simultaneous interaction with amino acid residues of the Ig2 and Ig3 modules of the same NCAM molecule

and thereby mimicking and/or modulating the interaction between the Ig2-to-Ig2 and the Ig3-to-Ig1 modules of two individual NCAM molecules interacting through said modules,

vii) candidate compounds, which are capable of simultaneous interaction with amino acid residues of the Ig2 and Ig3 modules of the same NCAM molecule and thereby mimicking and/or modulating the interaction between the Ig2-to-Ig3 and the Ig3-to-Ig1 modules of two individual NCAM molecules interacting through said modules,

viii) candidate compounds, which are capable of simultaneous interaction with amino acid residues of the Ig2 and Ig3 modules of the same NCAM molecule and thereby mimicking and/or modulating the interaction between the Ig2-to-Ig3 and the Ig3-to-Ig2 modules of two individual NCAM molecules interacting through said modules.

The candidate compound may be any compound, which as capable of the above interactions. Such a compound may, for example, be selected from the group comprising combinatorial libraries of peptides, lipids, carbohydrates or other organic molecules, or co-polymers of amino acids with other organic compounds. In a preferred embodiment, the candidate compound of the invention is a peptide. Alternatively, the compound may an antibody molecule capable of selectively binding to an epitope located within or in close proximity to the binding site. The close proximity is defined herein as a distance of about 50 to 500 Å between an amino acid residue of the epitope and an amino acid residue of the binding site. An antibody molecule, which binds to a distant epitope (at a distance of more then 500 Å from an amino acid residue of the binding site), and which binding leads to a conformational change in the Ig1-2-3 module influencing thereby the homophylic binding through the binding site, is also concerned.

A preferred candidate compound is the compound comprising an amino acid sequence derived from the binding site described herein. By "derived" is meant that the amino acid sequence comprises a fragment of the amino sequence ig the Ig1-2-3 NCAM module, said fragment comprising the binding site or a part of the binding site, or the amino acid sequence comprises a sequence of amino acid residues which are homologeous to the residues involved in the interaction between two individual NCAM molecules through the binding site. Homology of the amino acid

residues may be about 60%, more preferred is about 70%, even more preferred is about 80%, such as for example 90%, and the most preferred is about 100%. The homology of one amino acid residue to another is defined as above.

Thus, a preferred candidate compound may comprise a sequence derived from the Ig1 module and/or Ig2 module and/or Ig3 module of NCAM. Non-limited examples of such compounds may be the compounds identified in the present application as SEQ ID NOs: 1-20, 40-43.

Thus, the present invention provides in one embodiment a compound having the amino acid sequence WFSPNGEKLSPNQ set forth in SEQ ID NO: 1.

In another embodiment a compound of the invention is having the amino acid sequence YKCVVTAEDGTQSE set forth in SEQ ID NO: 2.

In still another embodiment the invention provides a compound having the amino acid sequence TLVADADGFPEP set forth in SEQ ID NO: 3.

In yet another embodiment the invention provides a compound having the amino acid sequence QIRGIKKTD set forth in SEQ ID NO: 4.

In still yet another embodiment the invention provides a compound having the amino acid sequence DVR set forth in SEQ ID NO: 5.

Yet in another embodiment the compound of the invention is having the amino acid sequence RGIKKTD set forth in SEQ ID NO: 6.

In yet a further embodiment the invention provides a compound is having the amino acid sequence DVRRGIKKTD set forth in SEQ ID NO: 7.

In another aspect the invention concerns a compound is having the amino acid sequence KEGED set forth in SEQ ID NO: 8.

In yet another aspect a compound is having the amino acid sequence IRGIKKTD set forth in SEQ ID NO: 9.

The invention further provides a compound having the amino acid sequence KEGEDGIRGIKKTD set forth in SEQ ID NO: 10.

Moreover, in another embodiment the invention provides a compound having the amino acid sequence DKNDE set forth in SEQ ID NO: 11.

In still another embodiment the invention concerns a compound having the amino acid sequence TVQARNSIVNAT set forth in SEQ ID NO: 12.

In yet another embodiment of the invention the compound is having the amino acid sequence SIHLKVFAK set forth in SEQ ID NO: 13.

In yet another embodiment the compound is having the amino acid sequence LSNNYLQIR set forth in SEQ ID NO: 14.

In a further embodiment the invention provides a compound having the amino acid sequence RFIVLSNNYLQI set forth in SEQ ID NO: 15.

Further, in yet another embodiment the invention provides a compound having the amino acid sequence KKDVRFIVLSNNYLQI set forth in SEQ ID NO: 16..

Furthermore, in yet another embodiment the invention provides a compound having the amino acid sequence QEFKEGEDAVIV set forth in SEQ ID NO: 17.

The invention further provides a compound having the amino acid sequence KEGEDAVIVCD set forth in SEQ ID NO: 18.

In other embodiments the invention concerns the compounds having the amino acid sequences
GEISVGESKFFL (SEQ ID NO: 19)
KHIFSDDSSELTIRNVDKNDE (SEQ ID NO: 20),
AFSPNGEKLSPNQ (SEQ ID NO: 40),
AKSVVTAEDGTQSE (SEQ ID NO: 41),
DVRRGIKKTD (SEQ ID NO: 42) or

QIRGIKKTD (SEQ ID NO: 43).

The above sequences are also concerned as preferred candidate compounds of the invention. Fragments or variants of these sequences are also included in the scope of the invention as candidate compounds capable of the interaction(s) and/or effects of the original sequences, namely the sequences, which they are derived from or homologous.

The identified above sequences according to the invention represent different fragments a homophylic binding site of NCAM in the Ig1-2-3 module and are capable of modulation of differentiation and/or survival of an NCAM-presenting cell.

According to invention the sequences identified above may be used for the manufacture of a medicament for the treatment of a condition or disease wherein the modulation of NCAM homophtlic interaction would lead to improvement or rescue.

Additionally, the above sequences may be used for the production of an antibody capable of recognising and specifically binding to the binding site of the invention. Such an antibody according to the invention is capable of at least one of the biological activities of the compound described above and may in some embodiments be advantageously used for medical applications as the compound of the invention

Production of peptide sequences

The peptide compounds of the present invention may be prepared by any conventional synthetic methods, recombinant DNA technologies, enzymatic cleavage of full-length proteins which the peptide sequences are derived from, such as for example NCAM molecules of different species origin, or a combination of said methods.

<u>Recombinant preparation</u>

Recombinant preparation is a preferred method for the production of long chain polypeptide sequences, such as for example the Ig1-2-3 module, or individual modules of NCAM, such as Ig1, Ig2 or Ig3, or combinations thereof such as Ig1-2, Ig1-3 or Ig2-3. However, shorter peptide fragments of 15-50 amino acids long comprising the sequences derived from the binding site of the invention may also be prepared recombinantly using any of the below described technologies.

The DNA sequence encoding a peptide or the corresponding full-length protein the peptide originates from, or a fragment thereof, may be prepared synthetically by established standard methods, e.g. the phosphoamidine method described by Beaucage and Caruthers, 1981, Tetrahedron Lett. 22:1859-1869, or the method described by Matthes et al., 1984, EMBO J. 3:801-805. According to the phosphoamidine method, oligonucleotides are synthesised, e.g. in an automatic DNA synthesiser, purified, annealed, ligated and cloned in suitable vectors.

The DNA sequence encoding a peptide may also be prepared by fragmentation of the DNA sequences encoding the corresponding full-length protein, e. g. NCAM protein, using DNAase I according to a standard protocol (Sambrook et al., Molecular cloning: A Laboratory manual. 2 rd ed., CSHL Press, Cold Spring Harbor, NY, 1989). The DNA encoding a full-length protein may alternatively be fragmented using specific restriction endonucleases. The fragments of DNA are further purified using standard procedures described in Sambrook et al., Molecular cloning: A Laboratory manual. 2 rd ed., CSHL Press, Cold Spring Harbor, NY, 1989.

The DNA sequence encoding a full-length protein may also be of genomic or cDNA origin, for instance obtained by preparing a genomic or cDNA library and screening for DNA sequences coding for all or part of the full-length protein by hybridisation using synthetic oligonucleotide probes in accordance with standard techniques (cf. Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor, 1989). The DNA sequence may also be prepared by polymerase chain reaction using specific primers, for instance as described in US 4,683,202 or Saiki et al., 1988, Science 239:487-491.

The DNA sequence is then inserted into a recombinant expression vector, which may be any vector, which may conveniently be subjected to recombinant DNA

procedures. The choice of vector will often depend on the host cell into which it is to be introduced. Thus, the vector may be an autonomously replicating vector, i.e. a vector that exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g. a plasmid. Alternatively, the vector may be one which, when introduced into a host cell, is integrated into the host cell genome and replicated together with the chromosome(s) into which it has been integrated.

In the vector, the DNA sequence encoding a peptide or a full-length protein should be operably connected to a suitable promoter sequence. The promoter may be any DNA sequence, which shows transcriptional activity in the host cell of choice and may be derived from genes encoding proteins either homologous or heterologous to the host cell. Examples of suitable promoters for directing the transcription of the coding DNA sequence in mammalian cells are the SV 40 promoter (Subramani et al., 1981, Mol. Cell Biol. 1:854-864), the MT-1 (metallothionein gene) promoter (Palmiter et al., 1983, Science 222: 809-814) or the adenovirus 2 major late promoter. A suitable promoter for use in insect cells is the polyhedrin promoter (Vasuvedan et al., 1992, FEBS Lett. 311:7-11). Suitable promoters for use in yeast host cells include promoters from yeast glycolytic genes (Hitzeman et al., 1980, J. Biol. Chem. 255:12073-12080; Alber and Kawasaki, 1982, J. Mol. Appl. Gen. 1: 419-434) or alcohol dehydrogenase genes (Young et al., 1982, in Genetic Engineering of Microorganisms for Chemicals, Hollaender et al, eds., Plenum Press, New York), or the TPI1 (US 4,599,311) or ADH2-4c (Russell et al., 1983, Nature 304:652-654) promoters. Suitable promoters for use in filamentous fungus host cells are, for instance, the ADH3 promoter (McKnight et al., 1985, EMBO J. 4:2093-2099) or the tpiA promoter.

The coding DNA sequence may also be operably connected to a suitable terminator, such as the human growth hormone terminator (Palmiter et al., op. cit.) or (for fungal hosts) the TPI1 (Alber and Kawasaki, op. cit.) or ADH3 (McKnight et al., op. cit.) promoters. The vector may further comprise elements such as polyadenylation signals (e.g. from SV 40 or the adenovirus 5 Elb region), transcriptional enhancer sequences (e.g. the SV 40 enhancer) and translational enhancer sequences (e.g. the ones encoding adenovirus VA RNAs).

The recombinant expression vector may further comprise a DNA sequence enabling the vector to replicate in the host cell in question. An example of such a sequence (when the host cell is a mammalian cell) is the SV 40 origin of replication. The vector may also comprise a selectable marker, e.g. a gene the product of which complements a defect in the host cell, such as the gene coding for dihydrofolate reductase (DHFR) or one which confers resistance to a drug, e.g. neomycin, hydromycin or methotrexate.

The procedures used to ligate the DNA sequences coding the peptides or full-length proteins, the promoter and the terminator, respectively, and to insert them into suitable vectors containing the information necessary for replication, are well known to persons skilled in the art (cf., for instance, Sambrook et al., op.cit.).

To obtain recombinant peptides of the invention the coding DNA sequences may be usefully fused with a second peptide coding sequence and a protease cleavage site coding sequence, giving a DNA construct encoding the fusion protein, wherein the protease cleavage site coding sequence positioned between the HBP fragment and second peptide coding DNA, inserted into a recombinant expression vector, and expressed in recombinant host cells. In one embodiment, said second peptide selected from, but not limited by the group comprising glutathion-S-reductase, calf thymosin, bacterial thioredoxin or human ubiquitin natural or synthetic variants, or peptides thereof. In another embodiment, a peptide sequence comprising a protease cleavage site may be the Factor Xa, with the amino acid sequence *IEGR*, enterokinase, with the amino acid sequence *DDDDK*, thrombin, with the amino acid sequence *LVPR/GS*, or *Acharombacter lyticus,* with the amino acid sequence X*K*X, cleavage site.

The host cell into which the expression vector is introduced may be any cell which is capable of expression of the peptides or full-length proteins, and is preferably a eukaryotic cell, such as invertebrate (insect) cells or vertebrate cells, e.g. *Xenopus laevis* oocytes or mammalian cells, in particular insect and mammalian cells. Examples of suitable mammalian cell lines are the HEK293 (ATCC CRL-1573), COS (ATCC CRL-1650), BHK (ATCC CRL-1632, ATCC CCL-10) or CHO (ATCC CCL-61) cell lines. Methods of transfecting mammalian cells and expressing DNA sequences introduced in the cells are described in e.g. Kaufman and Sharp, J. Mol.

Biol. 159, 1982, pp. 601-621; Southern and Berg, 1982, J. Mol. Appl. Genet. 1:327-341; Loyter et al., 1982, Proc. Natl. Acad. Sci. USA 79: 422-426; Wigler et al., 1978, Cell 14:725; Corsaro and Pearson, 1981, in Somatic Cell Genetics 7, p. 603; Graham and van der Eb, 1973, Virol. 52:456; and Neumann et al., 1982, EMBO J. 1:841-845.

Alternatively, fungal cells (including yeast cells) may be used as host cells. Examples of suitable yeast cells include cells of *Saccharomyces spp.* or *Schizosaccharomyces spp.*, in particular strains of *Saccharomyces cerevisiae*. Examples of other fungal cells are cells of filamentous fungi, e.g. *Aspergillus spp.* or *Neurospora spp.*, in particular strains of *Aspergillus oryzae* or *Aspergillus niger*. The use of *Aspergillus spp.* for the expression of proteins is described in, e.g., EP 238 023.

The medium used to culture the cells may be any conventional medium suitable for growing mammalian cells, such as a serum-containing or serum-free medium containing appropriate supplements, or a suitable medium for growing insect, yeast or fungal cells. Suitable media are available from commercial suppliers or may be prepared according to published recipes (e.g. in catalogues of the American Type Culture Collection).

The peptides or full-length proteins recombinantly produced by the cells may then be recovered from the culture medium by conventional procedures including separating the host cells from the medium by centrifugation or filtration, precipitating the proteinaceous components of the supernatant or filtrate by means of a salt, e.g. ammonium sulphate, purification by a variety of chromatographic procedures, e.g. HPLC, ion exchange chromatography, affinity chromatography, or the like.

Synthetic preparation

Synthetic preparation is preferred when the short sequences of 3 to 50 amino acids are concerned.

The methods for synthetic production of peptides are well known in the art. Detailed descriptions as well as practical advice for producing synthetic peptides may be

found in Synthetic Peptides: A User's Guide (Advances in Molecular Biology), Grant G. A. ed., Oxford University Press, 2002, or in: Pharmaceutical Formulation: Development of Peptides and Proteins, Frokjaer and Hovgaard eds., Taylor and Francis, 1999.

Peptides may for example be synthesised by using Fmoc chemistry and with Acm-protected cysteins. After purification by reversed phase HPLC, peptides may be further processed to obtain for example cyclic or C- or N-terminal modified isoforms. The methods for cyclization and terminal modification are well-known in the art and described in detail in the above-cited manuals.

In a preferred embodiment the peptide sequences of the invention are produced synthetically, in particular, by the Sequence Assisted Peptide Synthesis (SAPS) method. Peptides may be synthesised either batchwise in a polyethylene vessel equipped with a polypropylene filter for filtration or in the continuous-flow version of the polyamide solid-phase method (Dryland, A. and Sheppard, R.C., (1986) J.Chem. Soc. Perkin Trans. I, 125 - 137.) On a fully automated peptide synthesiser (Cameron ET al., (1987), J. Chem. Soc. Chem.Commun, 270-272) using 9-fluorenylmethyloxycarbonyl (Fmoc) or tert. -Butyloxycarbonyl, (Boc) as N-a-amino protecting group and suitable common protection groups for side-chain functionalities.

Pharmaceutical composition

Once the candidate compound(s) of the invention has been identified it is further within the scope of the invention to provide a pharmaceutical composition comprising one or more compound(s). In the present context the term pharmaceutical composition is used synonymously with the term medicament.

The invention is further related to a pharmaceutical composition capable of preventing the death of NCAM presenting cells, promote cell differentiation of neural cells and neuronal plasticity, and stimulation of survival and regeneration of NCAM presenting cells and/or NCAM ligand presenting cells in several tissues and organs *in vivo* or *in vitro* as discussed herein, said composition comprising an effective amount of one or more of the compounds described above. The medicament of the

invention may comprise an effective amount of one or more of the compounds as defined above in combination with pharmaceutically acceptable additives. Such medicament may suitably be formulated for oral, percutaneous, intramuscular, intravenous, intracranial, intrathecal, intracerebroventricular, intranasal or pulmonal administration.

The present invention further concerns a medicament for the treatment of diseases and conditions of the central and peripheral nervous system, of the muscles or of various organs, wherein said medicament comprises an effective amount of one or more of the compounds as defined above or a composition as defined above in combination with pharmaceutically acceptable additives or carriers. Such medicament may suitably be formulated for oral, percutaneous, intramuscular, intravenous, intracranial, intrathecal, intracerebroventricular, intranasal or pulmonal administration.

Formulation

Strategies in formulation development of medicaments and compositions based on the compounds of the present invention generally correspond to formulation strategies for any other protein-based drug product. Potential problems and the guidance required to overcome these problems are dealt with in several textbooks, e.g. "Therapeutic Peptides and Protein Formulation. Processing and Delivery Systems", Ed. A.K. Banga, Technomic Publishing AG, Basel, 1995.

Injectables are usually prepared either as liquid solutions or suspensions, solid forms suitable for solution in, or suspension in, liquid prior to injection. The preparation may also be emulsified. The active ingredient is often mixed with excipients, which are pharmaceutically acceptable and compatible with the active ingredient. Suitable excipients are for example water, saline, dextrose, glycerol, ethanol or the like, and combinations thereof. In addition, if desired, the preparation may contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents, which enhance the effectiveness or transportation of the preparation.

Formulations of the compounds of the invention can be prepared by techniques known to the person skilled in the art. The formulations may contain

pharmaceutically acceptable carriers and excipients including microspheres, liposomes, microcapsules, nanoparticles or the like.

Administration

For most indications a localised or substantially localised application is preferred. The compounds are in particular used in combination with a prosthetic device such as a prosthetic nerve guide. Thus, in a further aspect, the present invention relates to a prosthetic nerve guide, characterised in that it comprises one or more of the compounds or the composition defined above. Nerve guides are known in the art.

The preparation may suitably be administered by injection, optionally at the site, where the active ingredient is to exert its effect. Additional formulations which are suitable for other modes of administration include suppositories, nasal, pulmonal and, in some cases, oral formulations. For suppositories, traditional binders and carriers include polyalkylene glycols or triglycerides. Such suppositories may be formed from mixtures containing the active ingredient(s) in the range of from 0.5% to 10%, preferably 1-2%. Oral formulations include such normally employed excipients as, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, and the like. These compositions take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders and generally contain 10-95% of the active ingredient(s), preferably 25-70%.

Other formulations are such suitable for nasal and pulmonal administration, e.g. inhalators and aerosols.

The active compound may be formulated as neutral or salt forms. Pharmaceutically acceptable salts include acid addition salts (formed with the free amino groups of the peptide compound) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic acid, oxalic acid, tartaric acid, mandelic acid, and the like. Salts formed with the free carboxyl group may also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine, and the like.

The preparations are administered in a manner compatible with the dosage formulation, and in such amount as will be therapeutically effective. The quantity to be administered depends on the subject to be treated, including, e.g. the weight and age of the subject, the disease to be treated and the stage of disease. Suitable dosage ranges are of the order of several hundred μg active ingredient per administration with a preferred range of from about 0.1 μg to 100 mg, such as in the range of from about 1 μg to 100 mg, and especially in the range of from about 10 μg to 50 mg. Administration may be performed once or may be followed by subsequent administrations. The dosage will also depend on the route of administration and will vary with the age and weight of the subject to be treated. A preferred dosis would be in the interval 0.5 mg to 50 mg per 70 kg body weight.

Some of the candidate compounds of the present invention are sufficiently active, but for others, the effect will be enhanced if the preparation further comprises pharmaceutically acceptable additives and/or carriers. Such additives and carriers will be known in the art. In some cases, it will be advantageous to include a compound, which promote delivery of the active substance to its target.

In another embodiment it may be advantageous to administer the candidate compound(s) according to the invention with other substances to obtain a synergistic effect. Examples of such other substances may be a growth factor, which can induce differentiation, or a hormone, or a transplant of cells, including a transplant of stem cells, or gene therapy, or immuno-therapy.

In many instances, it will be necessary to administrate the formulation multiple times. Administration may be a continuous infusion, such as intra-ventricular infusion or administration in more doses such as more times a day, daily, more times a week, or weekly. It is preferred that administration of the medicament is initiated before or shortly after the individual has been subjected to the factor(s) that may lead to cell death. Preferably the medicament is administered within 8 hours from the factor onset, such as within 5 hours from the factor onset. Many of the compounds exhibit a long-term effect whereby administration of the compounds may be conducted with long intervals, such as 1 week or 2 weeks.

In one embodiment of the invention the administration of the present compound may

be immediately after an acute injury, such as an acute stroke, or at the most 8 hours after said stroke in order for the present compound to have a stimulatory effect on cell survival. Further, in cases concerning proliferation and/or differentiation the administration according to the invention is not time dependent, i.e. it may be administered at any time.

Producing a pharmaceutical

In another aspect the invention relates to a process of producing a pharmaceutical composition, comprising mixing an effective amount of one or more of the compounds of the invention, or a pharmaceutical composition according to the invention with one or more pharmaceutically acceptable additives or carriers, and administer an effective amount of at least one of said compound, or said pharmaceutical composition to a subject.

In yet a further aspect the invention relates to a method of treating an individual suffering from one or more of the diseases discussed above by administering the said individual a compound as described herein or a pharmaceutical composition comprising said compound.

Medicament

The candidate compounds of the invention may advantageously be used for treating cells presenting NCAM molecules. NCAM is expressed in a variety of cells of the organism where it can serve as an adhesion molecule, receptor and/or as a ligand, depending on cellular type and/or cell environment. The authors of the present invention discovered that the candidate compounds that bind to different parts of the binding site described herein may have different effects on cellular adhesion, differentiation and/or survival and therefore said compounds are suggested herein for using as both direct stimulators of cell differentiation and survival, such as neural cell differentiation and survival, and as modulators of NCAM function in neural plasticity, such as for example synaptic plasticity. Thus, the candidate compounds of the invention may be used in the manufacture of medicaments to be used to treat various pathologic conditions of the peripheral and/or the central nervous system and/or muscles and other tissues expressing NCAM, such as trauma and/or disease, as well as conditions in which a fine modulation of NCAM function may be beneficial, such as stimulation of memory and learning.

Thus, a candidate compound of the invention may be for the manufacture of a medicament for treatment of normal, degenerated or damaged NCAM and/or NCAM ligand presenting cells.

In particular the compound and/or pharmaceutical composition of the invention may be used in the treatment of clinical conditions, such as Neoplasms such as malignant neoplasms, benign neoplasms, carcinoma in situ and neoplasms of uncertain behavior, diseases of endocrine glands, such as diabetes mellitus, psychoses, such as senile and presenile organic psychotic conditions, alcoholic psychoses, drug psychoses, transient organic psychotic conditions, Alzheimer's disease, cerebral lipidoses, epilepsy, general paresis [syphilis], hepatolenticular degeneration, Huntington's chorea, Jakob-Creutzfeldt disease, multiple sclerosis, Pick's disease of the brain, syphilis,Schizophrenic disorders, affective psychoses, neurotic disorders, personality disorders, including character neurosis, non-psychotic personality disorder associated with organic brain syndromes, paranoid personality disorder, fanatic personality, paranoid personality (disorder), paranoid traits, sexual deviations and disorders, mental retardation, disease in the nervous system and sense organs, cognitive anomalies, inflammatory disease of the central nervous system, such as meningitis, encephalitis, cerebral degenerations, such as Alzheimer's disease, Pick's disease, senile degeneration of brain, communicating hydrocephalus, obstructive hydrocephalus, Parkinson's disease including other extra pyramidal disease and abnormal movement disorders, spino-cerebellar disease, cerebellar ataxia, Marie's, Sanger-Brown, Dyssynergia cerebellaris myoclonica, primary cerebellar degeneration, such as spinal muscular atrophy, familial, juvenile, adult spinal muscular atrophy, motor neuron disease, amyotrophic lateral sclerosis, motor neuron disease, progressive bulbar palsy, pseudobulbar palsy, primary lateral sclerosis, other anterior horn cell diseases, anterior horn cell disease, unspecified, other diseases of spinal cord, syringomyelia and syringobulbia, vascular myelopathies, acute infarction of spinal cord (embolic) (nonembolic), arterial thrombosis of spinal cord, edema of spinal cord, subacute necrotic myelopathy, subacute combined degeneration of spinal cord in diseases classified elsewhere, myelopathy, drug-induced, radiation-induced myelitis, disorders of the autonomic nervous system, disorders of peripheral autonomic,

sympathetic, parasympathetic, or vegetative system, familial dysautonomia [Riley-Day syndrome], idiopathic peripheral autonomic neuropathy, carotid sinus syncope or syndrome, cervical sympathetic dystrophy or paralysis. peripheral autonomic neuropathy in disorders classified elsewhere, amyloidosis, diseases of the peripheral nerve system, brachial plexus lesions, cervical rib syndrome, costoclavicular syndrome, scalenus anterior syndrome, thoracic outlet syndrome, brachial neuritis or radiculitis, including in newborn; inflammatory and toxic neuropathy, including acute infective polyneuritis, Guillain-Barre syndrome, Postinfectious polyneuritis, polyneuropathy in collagen vascular disease, disorders affecting multiple structures of eye, purulent endophthalmitis, diseases of the ear and mastoid process, chronic rheumatic heart disease, ischaemic heart disease, arrhythmia, diseases in the pulmonary system, abnormality of organs and soft tissues in newborn, including in the nerve system, complications of the administration of anesthetic or other sedation in labor and delivery, diseases in the skin including Infection, insufficient circulation problem, injuries, including after surgery, crushing injury, burns. Injuries to nerves and spinal cord, including division of nerve, lesion in continuity (with or without open wound), traumatic neuroma (with or without open wound), traumatic transient paralysis (with or without open wound), accidental puncture or laceration during medical procedure, injury to optic nerve and pathways, optic nerve injury, second cranial nerve, injury to optic chiasm, injury to optic pathways, injury to visual cortex, unspecified blindness, injury to other cranial nerve(s), injury to other and unspecified nerves. Poisoning by drugs, medicinal and biological substances, genetic or traumatic atrophic muscle disorders; or for the treatment of diseases or conditions of various organs, such as degenerative conditions of the gonads, of the pancreas, such as diabetes mellitus type I and II, of the kidney, such as nephrosis.

Conditions of CNS/PNS

In another aspect of the invention the compounds are for the treatment of diseases or conditions of the central and peripheral nervous system, such as postoperative nerve damage, traumatic nerve damage, impaired myelination of nerve fibers, postischaemic damage, e.g. resulting from a stroke, Parkinson's disease, Alzheimer's disease, Huntington's disease, dementias such as multiinfarct dementia, sclerosis, nerve degeneration associated with diabetes mellitus, disorders affecting the circadian clock or neuro-muscular transmission, and schizophrenia,

mood disorders, such as manic depression; for treatment of diseases or conditions of the muscles including conditions with impaired function of neuro-muscular connections, such as after organ transplantation, or such as genetic or traumatic atrophic muscle disorders; or for treatment of diseases or conditions of various organs, such as degenerative conditions of the gonads, of the pancreas such as diabetes mellitus type I and II, of the kidney such as nephrosis and of the heart and bowel, and for the treatment of postoperative nerve damage, traumatic nerve damage, impaired myelination of nerve fibers, postischaemic, e.g. resulting from a stroke, Parkinson's disease, Alzheimer's disease, dementias such as multiinfarct dementia, sclerosis, nerve degeneration associated with diabetes mellitus, disorders affecting the circadian clock or neuro-muscular transmission, and schizophrenia, mood disorders, such as manic depression.

Preventing cell death

Further, the candidate compounds according to the invention may be used for preventing cell death of cells being implanted or transplanted. This is particularly useful when using compounds having a long-term effect.

In another aspect of the invention the candidate compounds may be synthesised and secreted from implanted or injected gene manipulated cells.

Heart muscles

Furthermore, the candidate compound and/or pharmaceutical composition may be for preventing cell death of heart muscle cells, such as after acute myocardial infarction, or after angiogenesis. Furthermore, in one embodiment the compound and/or pharmaceutical composition is for the stimulation of the survival of heart muscle cells, such as survival after acute myocardial infarction. In another aspect the compound and/or pharmaceutical composition is for re-vascularisation, such as after injuries.

Memory

In another aspect the candidate compound and/or pharmaceutical composition is used for stimulation of the ability to learn and/or of the short and/or long-term memory.

Regeneration

In one aspect of the invention treatment by the use of the candidate compounds according to the invention is useful for the stimulation of regenerating cells which are degenerating or at risk of dying due to a variety of factors, such as traumas and injuries, acute diseases, chronic diseases and/or disorders, in particular degenerative diseases normally leading to cell death, other external factors, such as medical and/or surgical treatments and/or diagnostic methods that may cause formation of free radicals or otherwise have cytotoxic effects, such as X-rays and chemotherapy.

Prion disease

The candidate compound or a pharmaceutical composition comprising thereof may also be used for treating the prion diseases. NCAM has been shown to be a molecular interaction partner with the cellular prion protein.

For wound-healing

It is also within the scope of the invention to use the candidate compound and/or pharmaceutical composition for the promotion of wound-healing. The present compounds are capable of interfering with cell adhesion and thereby promote the wound healing process.

Cancer

The invention further discloses the use of the candidate compound and/or pharmaceutical composition in the treatment of cancer. NCAM regulates motility and inhibits cancer cells from spreading.

References

Atkins, A.R., Osborne, M.J., Lashuel, H.A., Edelman, G.M., Wright, P.E., Cunningham, B.A., and Dyson, H.J. (1999). Association between the first two immunoglobulin-like domains of the neural cell adhesion molecule N-CAM. FEBS Lett. *451*, 162-168.

Atkins, A.R, Chung, J., Deechongkit, S., Little, E.B., Edelman, G.M., Wright, P.E., Cunningham, B.A., and Dyson, H.J. (2001). Solution structure of the third immunoglobulin domain of the neural cell adhesion molecule N-CAM: can solution studies define the mechanism of homophilic binding? J. Mol. Biol. *311*, 161-172.

Becker, J.W., Erickson, H.P., Hoffman, S., Cunningham, B.A., and Edelman, G.M. (1989). Topology of cell adhesion molecules. Proc. Natl. Acad. Sci. USA *86*, 1088-1092.

Berezin, V., Bock, E., and Poulsen, F.M. (2000). The neural cell adhesion molecule. Curr. Opin. Drug Discovery Dev. *3*, 605-609.

Bork, P., Downing, A.K., Kieffer, B., and Campbell, I.D. (1996). Structure and distribution of modules in extracellular proteins. Q. Rev. Biophys. *29*, 119-167.

Brieher, W.M., Yap, A.S., and Gumbiner, B.M. (1996). Lateral dimerization is required for the homophilic binding activity of C-cadherin. J Cell Biol. *135*, 487-496.

Brünger, A.T., Adams, P.A., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J-S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., and Warren, G.L. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Cryst. *D54*, 905-921.

Casasnovas, J.M., Stehle, T., Liu, J.H., Wang, J.H., and Springer, T.A. (1998). A dimeric crystal structure for the N-terminal two domains of intercellular adhesion molecule-1. Proc. Natl. Acad. Sci. USA. *95*, 4134-4139.

Chothia, C., and Jones, E.Y. (1997). The molecular structure of cell adhesion molecules. Annu. Rev. Biochem. *66*, 823-862.

Cole, G.J., and Akeson, R. (1989). Identification of a heparin binding domain of the neural cell adhesion molecule N-CAM using synthetic peptides. Neuron *2*, 1157-1165.

Collaborative Computational Project, number 4. (1994). The CCP4 Suite: Programs for Protein Crystallography. Acta Cryst. *D50*, 760-763.

Covault, J., and Sanes, J.R. (1985). Neural cell adhesion molecule (NCAM) accumulates in denervated and paralyzed skeletal muscles. Proc. Natl. Acad. Sci. USA *82*, 4544-4548.

Conte, L.L., Chothia, C., and Janin, J. (1999). The atomic structure of protein-protein recognition sites. J. Mol. Biol. *285*, 2177-98.

Cremer, H., Lange, R., Christoph, A., Plomann, M., Vopper, G., Roes, J., Brown, R., Baldwin, S., Kraemer, P., Scheff, S., Barthels, D., Rajewsky, K., and Wille, W. (1994). Inactivation of the N-CAM gene in mice results in size reduction of the olfactory bulb and deficits in spatial learning. Nature *367*, 455-459.

Cunningham, B.A., Hemperly, J.J., Murray, B.A., Prediger, E.A., Brackenbury, R., and Edelman, G.M. (1987). Neural cell adhesion molecule: Structure, immunoglobulin-like domains, cell surface modulation, and alternative RNA splicing. Science *236*, 799-806.

Drejer J. and Schousboe A. (1989) Selection of a pure cerebellar granule cell culture by kainate treatment. Neurochem Res. *14*:751-4

Edelman, G.M., and Crossin, K.L. (1991). Cell adhesion molecules: implications for a molecular histology. Annu. Rev. Biochem. *60*, 155-190.

Eksterowicz JE, Evensen E, Lemmen C, Brady GP, Lanctot JK, Bradley EK, Saiah E, Robinson LA, Grootenhuis PD, Blaney JM. (2002) Coupling structure-based

design with combinatorial chemistry: application of active site derived pharmacophores with informative library design. J Mol Graph Model. *20,* 469-77.

Flocco, M.M., and Mowbray, S.L. (1994). Planar stacking interactions of arginine and aromatic side-chains in proteins. J. Mol. Biol. *235*, 709-717.

Freigang, J., Proba, K., Leder, L., Diederichs, K., Sonderegger, P., and Welte, W. (2000). The crystal structure of the ligand binding module of axonin-1/TAG-1 suggests a zipper mechanism for neural cell adhesion. Cell *101*, 425-433.

Gunning, P., Leavitt, J., Muscat, G., Ng, S.Y., and Kedes, L. (1987). A human beta-actin expression vector system directs high-level accumulation of antisense transcripts. Proc. Natl. Acad. USA. *84*, 4831-4835.

Hall, A.K., and Rutishauser, U. (1987). Visualization of neural cell adhesion molecule by electron microscopy. J. Cell Biol. *104*, 1579-1586.

Hunter, I., Sawa, H., Edlund, M., and Öbrink, B. (1996). Evidence for regulated dimerization of cell-cell adhesion molecule (C-CAM) in epithelial cells. Biochem. J. *320*, 847-853.

Janin, J. (1997). Specific versus non-specific contacts in protein crystals. Nature Struct. Biol. *4*, 973-974.

Jensen, P.H., Soroka, V., Thomsen, N.K., Ralets, I., Berezin, V., Bock, E., and Poulsen, F.M. (1999). Structure and interactions of NCAM modules 1 and 2 - basic elements in neural cell adhesion. Nature Struct. Biol. *6*, 486-493.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. *A47*, 110-119.

Jones, E.Y., Davis, S.J., Williams, A.F., Harlos, K., and Stuart, D.I. (1992). Crystal structure at 2.8 A resolution of a soluble form of the cell adhesion molecule CD2. Nature *360,* 232-239.

Jones, S., and Thornton, J.M. (1996). Principles of protein-protein interactions. Proc. Natl. Acad. Sci. USA *93*, 13-20.

Jørgensen, O.S., and Bock, E. (1974). Brain-specific synaptosomal membrane proteins demonstrated by crossed immunoelectrophoresis. J. Neurochem. *23*, 879-880.

Kallapur, S.G., and Akeson, R.A. (1992). The neural cell adhesion molecule (NCAM) heparin binding domain binds to cell surface heparan sulfate proteoglycans. J. Neurosci. Res. *33*, 538-548.

Kasper, C., Stahlhut, M., Berezin, V., Maar, T.E., Edvardsen, K., Kiselyov, V.V., Soroka, V., and Bock, E. (1996). Functional characterization of NCAM fibronectin type III domains: demonstration of modulatory effects of the proline-rich sequence encoded by alternatively spliced exons a and AAG. J. Neurosci. Res. *46*, 173-186.

Kasper, C., Rasmussen, H., Kastrup, J.S., Ikemizu, S., Jones, E.Y., Berezin, V., Bock, E., and Larsen, I.K. (2000). Structural basis of cell-cell adhesion by NCAM. Nature Struct. Biol. *7*, 389-393.

Kiselyov, V.V., Berezin, V., Maar, T., Soroka, V., Edvardsen, K., Schousboe, A., and Bock, E. (1997). The first Ig-like NCAM domain is involved in both double reciprocal interaction with the second Ig-like NCAM domain and in heparin binding. J.Biol.Chem. *272*, 10125-10134.

Kleywegt, G.J., and Jones, T.A. (1996). Phi/psi-chology: Ramachandran revisited. Structure *4*, 1395-1400.

Kolkova, K., Novitskaya, V., Pedersen, N., Berezin, V., and Bock, E. (2000). Neural cell adhesion molecule-stimulated neurite outgrowth depends on activation of protein kinase C and the Ras-mitogen-activated protein kinase pathway. J. Neurosci. *20*, 2238-2246.

Kostrewa, D., Brockhaus, M., D'Arcy, A., Dale, G.E., Nelboeck, P., Schmid, G., Mueller, F., Bazzoni, G., Dejana, E., Bartfai, T., Winkler, F.K., and Hennig, M. (2001). X-ray structure of junctional adhesion molecule: structural basis for homophilic adhesion via a novel dimerization motif. EMBO J. *20*, 4391-4398.

Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Cryst. *24*, 946-950.

Kristiansen, L.V., Marques, F.A., Soroka, V., Rønn, L.C., Kiselyov, V., Pedersen, N., Berezin, V., and Bock E. (1999). Homophilic NCAM interactions interfere with L1 stimulated neurite outgrowth. FEBS Lett. *464*, 30-34.

Leahy, D.J., Aukhil, I., and Erickson, H.P. (1996). 2.0 A crystal structure of a four-domain segment of human fibronectin encompassing the RGD loop and synergy region. Cell *84*, 155-164.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. *26*, 283-291.

Merritt, E.A. and Bacon, D.J. (1997). Raster3D Photorealistic Molecular Graphics. Methods Enzymol. *277*, 505-524.

Milev, P., Friedlander, D.R., Sakurai, T., Karthikeyan, L., Flad, M., Margolis, R.K., Grummet, M., and Margolis, R.U. (1994). Interactions of the chondroitin sulfate proteoglican phosphacan, the extracellular domain of a receptor-type protein tyrosine phosphatase, with neurons, glia, and neural cell adhesion molecules. J. Cell Biol. *121*, 1409-1421.

Miyahara, M., Nakanishi, H., Takahashi, K., Satoh-Horikawa, K., Tachibana, K., and Takai, Y. (2000). Interaction of nectin with afadin is necessary for its clustering at cell-cell contact sites but not for its cis dimerization or trans interactions. J. Biol. Chem. *275*, 613-618.

Muller, D., Wang, C., Skibo, G., Toni, N., Cremer, H., Calaora, V., Rougon, G., and Kiss, J.Z. (1996). PSA-NCAM is required for activity-induced synaptic plasticity. Neuron *3*, 413-422.

Navaza, J., and Saludjian, P. (1997). AmoRe: An automated molecular replacement program package. Methods Enzymol. *276*, 581-594.

Nybroe, O., Moran, N., and Bock, E. (1989). Equilibrium binding analysis of neural cell adhesion molecule binding to heparin. J. Neurochem. *52*, 1947-1949.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307-326.

Perrakis, A., Morris, R., and Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. Nature Struct. Biol. *6*, 458-463.

Probstmeier, R., Kuhn, K., and Schachner, M. (1989). Binding properties of the neural cell adhesion molecule to different components of the extracellular matrix. J. Neurochem. *53*, 1794-1801.

Ranheim, T.S., Edelman, G.M., and Cunningham, B.A. (1996). Homophilic adhesion mediated by the neural cell adhesion molecule involves multiple immumoglobulin domains. Proc. Natl. Acad. Sci. USA. *93*, 4071-4075.

Rao, Y., Wu, X-F., Gariepy J., Rutishauser, Urs., and Siu, C-H. (1992). Identification of a peptide sequence involved in homophilic binding in the neural cell adhesion molecule NCAM. J. Cell Biol. *118*, 937-949.

Rao, Y., Zhao, X., and Siu, C.H. (1994). Mechanisms of homophilic binding mediated by the neural cell adhesion molecule NCAM. Evidence for isologous interaction. J. Biol. Chem. *269*, 27540-275448.

Rønn, L.C., Ralets, I., Hartz, B.P., Bech, M., Berezin, A., Berezin, V., Moller, A., and Bock, E. (2000). A simple procedure for quantification of neurite outgrowth based on stereological principles. J. Neurosci. Meth. *100*, 25-32.

Sandig, M., Rao, Y., and Siu, C-H. (1994). The homophilic binding site of the neural cell adhesion molecule NCAM is directly involved in promoting neurite outgrowth from cultured neural retinal cells. J. Biol. Chem. *269*, 14841-14848.

Shapiro, L., Fannon, A.M., Kwong, P.D., Thompson, A., Lehmann, M.S., Grubel, G., Legrand, J-F., Als-Nielsen, J., Colman, D.R., and Hendrickson, W.A. (1995). Structural basis of cell-cell adhesion by cadherins. Nature *374*, 327-337.

Soroka, V., Kiryushko, D., Novitskaya, V., Rønn, L.C., Poulsen, F.M., Holm, A., Bock, E., and Berezin, V. (2002). Induction of neuronal differentiation by a peptide corresponding to the homophilic binding site of the second Ig module of NCAM. J. Biol. Chem. 227, 24676-24683.

Su, X-D., Gastinel, L.N., Vaughn, D.E., Faye, I., Poon, P. and Bjorkman P.J. (1998). Crystal structure of hemolin: A horseshoe shape with implications for homophilic adhesion. Science *281*, 991-995.

Takeda, H., Shimoyama, Y., Nagafuchi, A., and Hirohashi, S. (1999). E-cadherin functions as a cis-dimer at the cell-cell adhesive interface in vivo. Nature Struct. Biol. *6*, 310-312.

Tomasiewicz, H., Ono, K., Yee, D., Thompson, C., Goridis, C., Rutishauser, U., and Magnuson, T. (1993). Genetic deletion of a neural cell adhesion molecule variant (NCAM-180) produces defects in the central nervous system. Neuron *11*, 1163-1174.

Thomsen, N.K., Soroka, V., Jensen, P.H., Berezin, V., Bock, E., and Poulsen, F.M. (1996). The three-dimensional structure of the first domain of neural cell adhesion molecule. Nature Struct. Biol. *3*, 581-585.

Wu, Y.Y., and Bradshaw, R.A. (1995). PC12-E2 cells: a stable variant with altered responses to growth factor stimulation. J. Cell. Physiol. *164*, 522-532.

Wu, H., Kwong, P.D., and Hendrickson, W.A. (1997). Dimeric association and segmental variability in the structure of human CD4. Nature *387*, 527-530.

Experimentals

The following is a non-limiting examples of the production of the candidate compounds of the invention, a fragment of NCAM comprising the Ig1-2-3 module and fragments thereof, such as Ig1, Ig2, Ig3, or Ig1-2, or Ig2-3, description of the crystalline protein comprising the Ig1-2-3 module, and the biological testing of selected candidate compounds.

Production of the Ig1-2-3 and Ig3 fragments of NCAM

The NCAM Ig1-2-3 and Ig3 fragments were produced as recombinant proteins in the yeast *P. pastoris* expression system (Invitrogen). The cDNA fragments encoding Ig1-2-3 and Ig3 of rat NCAM (NCBI accession number NP_113709), corresponding to residues 1-289 and 191-289, respectively, were synthesized by PCR using rat NCAM cDNA as a template. The following DNA primers were used for cloning of Ig1-2-3 and Ig3, respectively: upper (5'-TCT CTC GAG TTC TGC AGG TAG ATA TTG TT-3') (SEQ ID NO: 37) and lower (5'-AAA CCC GGG TTA CTT TGC AAA GAC CTT-3') (SEQ ID NO: 30), upper (5'-GAA TAC GTA ACT GTC CAG GCC AGA C-3') (SEQ ID NO: 31) and lower (5'-AAA CCT AGG TTA CTT TGC AAA GAC CTT G-3') (SEQ ID NO: 32). The amplified cDNA fragments were subcloned into the pHIL-S1 and the pPIC9K plasmids (Invitrogen), respectively. The recombinant plasmids were linearized with the NsiI and SacI restriction enzymes, respectively, and used for transformation of the *P. pastoris* strain His 4 GS-115 (Invitrogen). Large-scale production of the recombinant proteins was performed employing a high-density feed-batch fermentation technique in a Biostat B fermentor (B. Braun Biotech Int. GmbH). Ig1-2-3 and Ig3 were purified from concentrated and desalted medium by anion-exchange chromatography on a HiTrap Q-Sepharose 5 ml column (Pharmacia), followed by gel filtration chromatography on a HiLoad 16/60 Superdex-75 column (Pharmacia). The Ig1-2-3 was enzymatically deglycosylated with PNGase-F endo-N-glycosidase (New England Biolabs) at 37 °C in PBS buffer pH 7.4. The authenticity of the protein fragments was confirmed by DNA sequencing of the recombinant plasmids, by amino acid sequencing of the 10-12 N-terminal residues, and by MALDI-TOF MS. The recombinant Ig1-2-3 and Ig3 fragments contained respectively two (RV) and five (EAEAY) additional N-terminal residues from the cloning vector. The purity of the proteins was at least 95% as estimated by SDS-PAGE.

Production of the Ig1-2-3 and Ig3 mutants

An Ig1-2-3 mutant (Ig1-2-3mut) containing the substitutions E11A, E16A, and K18A was produced as a recombinant protein in the yeast *P. pastoris* expression system following the procedure described for the Ig1-2-3 fragment. The three mutations were introduced by PCR using the following DNA primer: upper (5'-CTG CAG GTA GAT ATT GTT CCC AGC CAA GGA GCC ATC AGC GTT GGA GCC TCC GCC TTC TTC CTG TGT CAA GTG GCA-3') (SEQ ID NO: 33).

Two Ig3 mutants containing the substitutions: R198A, D249G, E287A (Ig3mut1) and K285A, F287A (Ig3mut2) were produced as recombinant proteins in the yeast *P. pastoris* expression system following the procedure described for the Ig3 fragment. Mutations were introduced by PCR using the following DNA primers: upper1 (5'-AAA TAC GTA ACT GTC CAG GCC GCC CAG AGC ATC GTG-3') (SEQ ID NO: 38), upper2 (5'-GGC GAC AGT TCG GCG TTA ACC ATC AGG AAT GTG GAC-3') (SEQ ID NO: 34), and lower (5'-GGT TAA CGC CGA ACT GTC GCC ACT GAA GAT GTG CTT CTC-3') (SEQ ID NO: 35) for Ig3mut1, and lower (5'-AAA CTT AGG TTA CTT TGC TGC GAC TGC GAG GTG GAT GGA GGC ATC-3') (SEQ ID NO: 36) for Ig3mut2. The DNA constructs of Ig1-2-3mut, Ig3mut1, and Ig3mut2 were verified by DNA sequencing. Folding of the Ig3 module and its mutants, as well as presence of carbohydrates, was confirmed by one-dimensional proton NMR spectra recorded at 800 MHz on a Varian NMR spectrometer (Varian Inc.) at 25°C in PBS buffer pH 7.4.

Preparation of peptides

Peptides were synthesized using the 9-fluorenylmethoxycarbonyl (Fmoc) protection strategy on a TentaGel resin (Rapp Polymere) using Fmoc protected amino acids (Calbiochem-Novabiochem). Peptides were at least 85% pure as estimated by MALDI-TOF MS. All peptides were synthesized with free NH_2 and carboxy-amidated COOH groups.

Crystallization and data collection

Crystals of NCAM Ig1-2-3 were grown at 18°C using the hanging-drop vapor diffusion method, with drops of equal volumes of reservoir and protein solutions (4 mg ml^{-1} in 5 mM Na phosphate, 150 mM NaCl, pH 7.4). The reservoir solution contained 14-17% w/v PEG 4000, 450 mM Li sulfate, 100 mM Na acetate, pH 5.2. The crystals belong to space group $I2_12_12_1$ with one molecule in the asymmetric unit

and cell dimensions of a = 51.5, b = 108.5, and c = 149.0 Å. The crystals were flash cooled in liquid nitrogen using 15% v/v glycerol as cryoprotectant. Two data sets were collected on the same crystal. The high-resolution data were collected to 2.0 Å at 120 K at beamline I711, Max-Lab, Lund, Sweden, and the low-resolution data were collected to 3.5 Å at 120 K on a Rigaku RU300 rotating anode equipped with a MAR345 image plate detector. The data sets were combined and processed with DENZO/SCALEPACK (Otwinowski and Minor, 1997) and the CCP4 suite of programs (Collaborative Computational Project No. 4, 1994).

Structure determination and refinement

The structure was determined by molecular replacement with the programs AmoRe (Navaza and Saludjan, 1997) and CNS version 1.0 (Brünger et al., 1998), using the X-ray structures of the Ig2 and Ig1 modules of NCAM (Kasper et al., 2000) as search models. Initially, the position of the Ig2 module was located using AmoRe. The Ig1 module was subsequently located using CNS. An electron density map was calculated based on phase information from Ig1 and Ig2. Residues of Ig3 were gradually built into this map. Map interpretation and model building were carried out using the program O (Jones et al., 1991). After several building and refinement cycles, ARP/wARP version 5.1 (Perrakis et al., 1999) was used to rebuild 233 out of 291 residues of NCAM Ig1-2-3. CNS was used to carry out the final rounds of refinements. The final model contains amino acids (-1)-238 and 241-289, and 266 water molecules. Amino acids are numbered according to the mature sequence of NCAM. Residues Arg and Val originating from the cloning site were given negative integers -2 and -1, respectively. Using all reflections in the resolution range 50-2.0 Å, the R_{cryst} is 21.8% and the R_{free} is 23.8% (3% test set, corresponding to 828 reflections). Data collection and refinement statistics are given in Table 1 (Figure 1). Interdomain geometry was determined according to Bork et al. (1996), and buried accessible surface areas were calculated using the Protein-Protein Interaction Server (http://www.biochem.ucl.ac.uk/bsm/PP/server) (Jones and Thornton, 1996). Figures were prepared with the programs MOLSCRIPT, RASTER3D (Kraulis, 1991; Merritt and Bacon, 1997), and Insight II (Accelrys).

The atomic coordinates of the structure is demonstrated in the Table 2 (Figure 2).

Protein Data Bank ID code

The coordinates of the structure have been deposited with the Protein Data Bank under ID code 1QZ1.

Cell culture and immunostaining

The NCAM-expressing pheochromocytoma PC12-E2 cell line (Wu and Bradshaw, 1995) was a gift from Dr. Klaus Seedorf, Hagedorn Research Institute, Denmark. The cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 5% v/v fetal calf serum (FCS) and 10% v/v horse serum (HS), 100 U ml^{-1} penicillin, 100 μg ml^{-1} streptomycin (all from Gibco BRL) at 37°C in a humidified atmosphere containing 5% CO_2.

The fibroblastoid mouse cell line, L929 (European Cell Culture Collection), was stably transfected with the eukaryotic expression vector pHβ-Apr-1-neo (Gunning et al., 1987) containing a full-length cDNA encoding human 140 kDa NCAM-B or the vector alone. The NCAM cDNA did not contain the exons VASE, a, b, c, or AAG. The cells were routinely grown at 37°C, 5% CO_2 in DMEM supplemented with 10% v/v FCS, 100 U ml^{-1} penicillin, and 100 μg ml^{-1} streptomycin. For analysis of neurite outgrowth, PC12-E2 cells (8,000 cells per well) were seeded on top of a confluent monolayer of transfected fibroblastoid L929 cells in four-well LabTek Tissue Culture Chamber Slides (NUNC). The cells were grown for 24 h in DMEM supplemented with 1% v/v HS, before analysis.

Cerebellar granule neurons (CGN) were prepared from Wistar rat pups of postnatal day 3. Cerebellar tissue was dissected in modified Krebs-Ringer solution kept on ice, and treated as described for the hippocampal neurons above. All cell cultures were incubated at 37°C in a humidified atmosphere containing 5 % CO_2. All animals were handled in accordance with the national guidelines for animal welfare. Primary cultures of CGN were plated at a density of 100,000 cells/cm^2 on poly-L-lysine coated 8-well permanox slides in Neurobasal-A medium (Gibco, BRL) supplemented with 2 % (v/v) B27, 0.5 % (v/v) glutamax, 100 U/ml penicillin, 100 μg/ml streptomycin and KCl, making the final concentration of KCl in the medium 40 mM. 24 hours after plating, cytosine-β-D-arabinofuranoside (Ara-C; Sigma-Aldrich) was added to a final concentration of 10 μM to avoid proliferation of glial cells, after which the neurons were allowed to differentiate for further six days at 37°C.

The glycosylated recombinant rat Ig3 module of NCAM (wildtype and mutated forms) or selected peptides were added immediately after seeding of cells in order to evaluate their inhibitory effects on adhesion, as reflected by interference with NCAM-mediated

neurite outgrowth. Ig3wt, Ig3mut1, and Ig3mut2 were tested at a concentration of 500 μg ml^{-1}. Proper controls were included and the person performing the experiments did not know the identity of the mutants or peptides.

To evaluate the length of processes of PC12–E2 cells, the co-cultures were fixed in 4% w/v paraformaldehyde for 25 min. After washing in PBS, cells were blocked with 10% v/v goat serum (DAKO) for 30 min and subsequently incubated for 1 h at room temperature with a mouse monoclonal anti-Thy-1 antibody (Caltag Laboratories) (1:100 in PBS containing 10% v/v goat serum). After washing, cells were incubated for 1 h at room temperature with Alexa-Fluor 568™ goat anti-mouse IgG (Molecular Probes) (1:1000 in PBS containing 10% goat serum). All washes were performed for 10 min in PBS, and repeated three times.

To evaluate the length of neurites of CGN, the neurons were fixed after 24 hours in culture with 4 % (v/v) formaldehyde for 20 minutes and thereafter immunostained using primary rabbit antibodies against GAP-43 and Alexa Fluor secondary goat anti-rabbit antibodies. Images of at least 200 neurons for each group in each individual experiment were obtained systematically by using computer assisted fluorescence microscopy

The total neurite length per cell was analyzed using the software ProcessLength (Rønn et al., 2000). Five independent experiments with the Ig3 module, its mutants, and the individual peptides were performed. In each experiment neurites from 200-300 cells were analyzed. In order to compare results of individual experiments and due to the inherently high variability of cell experiments, the data were normalized setting the difference between the average neurite length of PC12-E2 cells grown on NCAM-140-transfected and vector-transfected fibroblasts to 100%. Statistical evaluations were performed using a two-sided Student's *t*-test.

Dynamic light scattering (DLS) measurements

Measurements were performed using a DynaPro-MS/X instrument (Protein Solutions) at 18°C. The deglycosylated preparations of Ig1-2-3 (4 mg ml^{-1}), Ig1-2-3mut (4 mg ml^{-1}) and Ig3 (10 mg ml^{-1}) in PBS pH 7.4 were used to determine the molecular weight of the recombinant proteins in solution.

Results and Discussion

The X-ray structure of the Ig1-2-3 modules of NCAM

The X-ray structure of NCAM Ig1-2-3 was determined to 2.0 Å resolution (see Table 1 of Figure 1). In the structure of Ig1-2-3, the Ig1 and Ig2 modules are positioned in an extended conformation with Ig3 oriented at an angle of approximately 45° to the Ig1-Ig2 axis (Figure 3). The linker regions between Ig1-Ig2 and between Ig2-Ig3 are short and comprise only two (Lys98 – Leu99) and one (Asn190) residues, respectively. The overall structure of the Ig1 and Ig2 modules is very similar to the previously determined Ig1-2 structure (Kasper et al., 2000) with root mean square deviations (r.m.s.d.) of 0.7 (96 Cα atoms) and 0.8 Å (93 Cα atoms), respectively. In the Ig1-2-3 structure, the tilt angle between Ig1 and Ig2 is 11° and thereby differs by 13° compared to the Ig1-2 structure.

The 98-residue Ig3 module of rat NCAM adopts the topology of an intermediate type 1 (I1) set Ig module (Casasnovas et al., 1998). In the Ig3 module, the classical β-sandwich consists of two β-sheets with a total of nine β-strands (Figure 3B). The A, B, D, and E β-strands make up one sheet and the A', C, C', F, and G β-strands the second sheet. A cysteine bridge Cys216 – Cys269 connects the two β-sheets. All strands are anti-parallel except for the A' strand, which runs parallel to the C-terminal part of the G strand. Ig3 contains one site for N-linked glycosylation at Asn203 positioned in the A' strand. The E-F loop (residues Lys261 – Asp263) forms a 3_{10} α-helical turn. The overall structure of rat Ig3 is similar to the structure of chicken Ig3 (Atkins et al., 2001) with r.m.s.d. of 1.65 Å (95 Cα atoms).

Parallel interactions between Ig modules

Several characteristic interactions are observed in the structure of the NCAM Ig1-2-3 fragment which may be divided into two groups: Interactions where the long axes (N- to C-terminus) of two interacting Ig1-2-3 molecules are oriented in a parallel manner and interactions where the long axes are oriented in an anti-parallel manner. One parallel interaction and three major anti-parallel interactions are observed in the crystal.

The parallel, cross-like dimer interaction of NCAM Ig1-2-3 involves the Ig1 and Ig2 modules (Figure 5). The total buried surface area of this interface is 1594 $Å^2$ (per dimer), which is similar to that previously observed in the Ig1-2 cross-like dimers (Kasper et al., 2000). The most prominent feature of the Ig1-to-Ig2 interaction is the intercalation of two aromatic residues of Ig1, Phe19 and Tyr65, into hydrophobic pockets formed by Ig2 residues (Figure 5A), which was also observed in the Ig1-2 structure. However, a tighter Ig1 to Ig2 binding interface is observed in the Ig1-2-3

structure, where the hydroxyl group of Tyr65 forms a direct hydrogen bond (H-bond) with Glu171, instead of a water-mediated H-bond as observed in Ig1-2. Tyr65 also makes three H-bonds to the side chains of Lys133, Glu171, and Arg173. Arg173 forms part of the Ig2 hydrophobic pocket and makes two H-bonds to Thr63. The parallel orientation of the Arg173 and Phe19 side chains and the distance between the N□ı atom of the guanidinium group of Arg173 and the Cζ atom of the benzene ring of Phe19 (3.4 Å) suggest a cation-π interaction between these two residues (Flocco and Mowbray, 1994).

Dynamic Light Scattering (DLS) measurements showed that deglycosylated Ig1-2-3 forms a single species of molecules in solution with a molecular weight of ~78 kDa, corresponding to a dimer. In order to demonstrate that Ig1-2-3 dimerization is mediated by the observed Ig1 to Ig2 binding, we produced a mutant of Ig1-2-3 (Ig1-2-3mut) containing three Ala substitutions: E11A, E16A, and K18A. These mutations have previously been shown to completely abolish dimerization of the Ig1-2 NCAM fragment in solution (Jensen et al., 1999). In the present structure Glu11 and Glu16 form intramolecular salt bridges, respectively, with Arg177 and Lys98 from the Ig1 to Ig2 linker region (not shown). These salt bridges probably contribute to the proper orientation of Ig1 with respect to Ig2 and therefore are important for the Ig1-to-Ig2 interaction. Lys18 forms an H-bond with the carboxyl group of Arg177 from the Ig2 module stabilizing the Ig1-Ig2 interaction (Figure 5A). Lys18 is located near Phe19, which is the critical residue for the Ig1-to-Ig2 interaction as it was clearly demonstrated earlier (Atkins et al., 2001). Therefore, disruption of the Lys18 - Arg177 H-bond may affect the orientation of Phe19 leading to elimination of the Ig1-to-Ig2 interaction. The molecular weight of the Ig1-2-3mut fragment was determined by DLS to be ~34 kDa, indicating a monomer. This confirms that Ig1-2-3 dimerization is mediated by Ig1-to-Ig2 binding.

Parallel (*cis*) interactions are not uncommon among cell adhesion molecules. Thus, *cis* dimerization has been demonstrated for the cell adhesion molecules C-CAM1, C-CAM2, ICAM-1, nectin-2α, and JAM belonging to the Ig superfamily (Hunter et al., 1996; Casasnovas et al., 1998; Miyahara et al., 2000; Kostrewa et al., 2001) as well as for N-, E-, and C- cadherins (Shapiro et al., 1995; Takeda et al., 1999; Brieher et al., 1996). It was shown that the dimeric form of C-cadherin is capable of adhesion, whereas the monomeric form is not (Brieher et al., 1996).

Anti-parallel interactions between Ig modules

An anti-parallel interaction takes place between the Ig2 and Ig3 modules of two Ig1-2-3 molecules, thereby forming arrays of Ig1-2-3 dimers (Figure 4A,B). Ig2 of one molecule binds to Ig3 of a second molecule, and *vice versa* (Figure 3B). The residues involved are 112-115, 143-146, and 158-161 from the B-strand, CD-loop/D-strand, and E-strand of Ig2, and residues 200-205, 261, and 278-289 from the A'-strand, EF-loop, and G-strand of Ig3. A central element of this interaction is the intercalation of the side chain of Phe287 from Ig3 into a hydrophobic pocket formed by the side chains of Val145, Arg146, and Arg158 of the Ig2 module and Lys285 from Ig3. Arg158 is also involved in water-mediated hydrogen bonding to residues Lys261 and Ala288, and Gly159 makes a direct H-bond to Asn203.

The crystal packing leaves room for glycosylation at Asn203. In order to accommodate N-linked glycosylation at this site, the side chain of Asn203 has to adopt another rotamer conformation. Thereby, the carbohydrate will point away from the binding site and towards a solvent channel in the crystal, and consequently Asn203 will not interfere with Ig2-Ig3 interactions. An interaction between the two Ig3 modules is observed at the interface, as Gln196 makes a water-mediated H-bond with Gln278. The total buried surface of the Ig2-to-Ig3 interface is 1407 $Å^2$ per dimer. According to Janin (1997), the probability of finding a non-specific interface of the size of the Ig2-to-Ig3 contact is only 1.9%.

Another anti-parallel interaction between two Ig1-2-3 molecules is formed between two Ig2 modules (Figure 4C,D). This interaction involves residues 103-121 and 150-158 of the AA'-loop/A'-strand/A'B-loop and the DE-loop/E-strand and has the total buried surface of 958 $Å^2$ per dimer (Figure 4C). Here, the central residue appears to be Glu114, which makes two H-bonds to Ser151 (side chain and backbone). Apart from an extensive hydrogen-bonding network, especially through water molecules, Val117, Val119, Leu150, and Tyr154 of both Ig2 modules form a number of hydrophobic contacts with each other at the Ig2-to-Ig2 interface (not shown).

A slightly smaller anti-parallel interaction (858 $Å^2$ of total buried surface per dimer) is formed between the Ig1 and Ig3 modules (Figure 4C,D), involving residues 32-47 and 76-88 from the C-strand/CC'-loop/C'-strand/C'D-loop and F-strand/FG-loop/G-strand in Ig1, and residues 198, 213-223, and 248-253 from the A-strand, B-strand/BC-loop, and D-strand/DE-loop in Ig3 (Figure 5D). Arg198 and Asp249 form direct H-bonds to the backbone oxygen atoms of Ala81 and Glu82 and two salt bridges with Lys76, respectively. Additionally, one water-mediated H-bond is formed between Lys42 and Asp250, one between Ser44 and Gly220, and two between

Ser44 and Glu223. The conserved Phe36 and Phe221 are packed against Asp249 and Gln47, respectively. Together two Ig1-to-Ig3 interaction sites and one Ig2-to-Ig2 site make up a predominant contact between Ig1-2-3 dimers in the crystal (2654 Å^2) forming the second array of Ig1-2-3 dimers (Figure 4C,D) perpendicular to the Ig2-to-Ig3-mediated array (Figure 2A,B). Contact areas of similar sizes have been found in other CAMs. *Cis* dimers of human ICAM-1 and mouse JAM have 1100 Å^2 and 1200 Å^2 of total buried surface area (per dimer), respectively (Casasnovas et al. 1998; Kostrewa et al., 2001), whereas *trans* dimers of rat CD2 and chicken axonin-1/TAG-1 have even larger contact areas of 1300 Å^2 and 2000 Å^2 (Jones et al., 1992; Freigang et al., 2000).

Ig3 inhibits NCAM-dependent neurite outgrowth

NCAM-NCAM interaction is known to induce neurite outgrowth from NCAM-expressing PC12-E2 cells grown on a confluent monolayer of NCAM-expressing fibroblasts (Kolkova et al., 2000). Inhibition of the NCAM-NCAM interaction will therefore inhibit neurite outgrowth in PC12-E2 cells.

In order to examine the biological significance of the observed Ig1-to-Ig3 and Ig2-to-Ig3 contacts in the structure of NCAM Ig1-2-3, we tested the inhibitory effect of the recombinant Ig3 module on NCAM-NCAM adhesion. Furthermore, we prepared two Ig3 mutants containing mutations of the residues R198A, D249G, E253A (Ig3mut1) of the Ig1-to-Ig3 contact site (see Figure 5D) and K285A, F287A (Ig3mut2) of the Ig2-to-Ig3 contact site (see Figure 5B). In Figure 4 it can be seen that the wildtype Ig3 module (Ig3wt) indeed has an inhibitory effect, whereas both mutants are inactive, thereby strongly supporting that both the Ig1-to-Ig3 and Ig2-to-Ig3 contact sites are participating in homophilic interactions.

A similar co-culture test-system of NCAM-expressing chicken retinal ganglion cells grown on top of NCAM-140-transfected mouse L-cells has been successfully used to demonstrate a disruptive effect of mutations in the Ig3 module homophilic binding site (Ig1-to-Ig3 binding site in the present work) as well as to show an inhibition of neurite outgrowth by synthetic peptides representing this homophilic binding site (Sandig et al. 1994).

Interaction interface peptides inhibit neurite outgrowth

It has previously been demonstrated that peptides representing homophilic binding sequences from Ig3 and Ig2 modules of NCAM inhibit NCAM-mediated cell

aggregation (Rao et al., 1992; Sandig et al. 1994; Rao et al., 1994; Soroka et al. 2002). Therefore, in order to further examine the biological significance of the observed Ig1-to-Ig2, Ig1-to-Ig3, and Ig2-to-Ig3 contacts in the structure of NCAM Ig1-2-3, we tested the inhibitory effect of a series of peptides representing amino acid sequences from the observed contact areas (Figure 6,8-12).

The Ig1-to-Ig2 contact was represented by the P1-B peptide (10-GEISVGESKFFL-21) (SEQ ID NO: 19) that covers the B β-strand of Ig1 and containing the key residue Phe19 in the Ig1-to-Ig2 binding (Kasper et al., 2000; Atkins et al., 2001). As a negative control, two peptides GEISVGESKAFL (P1-B-F19A) (SEQ ID NO:21) and GEISVGESK*AA*L (P1-B-F19A-F20A) (SEQ ID NO: 22) containing a single Ala substitution of F19 and a double Ala substitution of both F19 and F20, respectively, were used.

The Ig1-to-Ig3 contact was represented by three peptides: AFSPNGEKLSPNQ (P1-CD) (SEQ ID NO: 40), AKSVVTAEDGTQSE (P1-FG) (SEQ ID NO: 41) and KHIFSDDSSELTIRNVDKNDE (P3-DE) (SEQ ID NO: 20). The P3-DE peptide covering the sequence of the D and E β-strands and the E-F loop of the Ig3 module is homologous to the sequence previously suggested to be a homophilic binding site in the Ig3 module of chicken NCAM (243-KYSFNYDGSELIIKKVDKSDE-263) (SEQ ID NO: 23) (Rao et al., 1992). As a negative control, a truncated version of the P3-DE peptide 244-KHIFSDDSSE-253 (P3-DE-trunc) (SEQ ID NO: 24) was used. The P3-DE-trunc peptide is homologous to the 243-KYSFNYDGSE-252 (SEQ ID NO: 25) chicken sequence which was less potent than the longer sequence (Rao et al., 1992).

The Ig2-to-Ig2 contact was represented by the peptide P2-A'B (QEFKEGEDAVIV (SEQ ID NO: 17).

The Ig2-to-Ig3 contact was represented by the peptides P2-CD (DVRRGIKKTD) (SEQ ID NO: 42) and P2-EF (QIRGIKKTD) (SEQ ID NO: 43) covering the sequences of CD- and EF- β-strands of the Ig2 module correspondingly, and P3-G (SIHLKVFAK) (SEQ ID NO: 13) from the Ig3 module. The sequence SIHLKVFAK (SEQ ID NO: 13) covers the C-terminal part of the G β-strand including the solvent-exposed Phe287. As negative controls, two peptides SIHLAVAAK (P3-G-K285A-F287S) (SEQ ID NO: 26) and SIHLAVGAK (P3-G-K285A-F287G) (SEQ ID NO: 27) with substitutions of K285 and F287 were used. Both P1-B and P3-G peptides contain two hydrophobic residues (Ile and Val/Leu) close to their N-termini and at least one Phe residue close to their C-termini. As a control peptide with similar

hydrophobic properties we selected a peptide 213-TLVADADGFPEP-224 (P3-B) (SEQ ID NO: 3) covering the B β-strand and B-C loop of the Ig3 module, and including Gly220, Phe221, and Glu223 involved in Ig1-to-Ig3 binding. In spite of sequence similarity with P1-B and P3-G peptides, the P3-B peptide was not active (Figure 6G). This is probably due to the fact that Phe221 in Ig3 is partially solvent exposed and Gly220 and Glu223 form water-mediated hydrogen bonds (Figure 5D). In contrast, the peptides P1-B, P3-DE, and P3-G either contain Phe buried in a hydrophobic pocket or residues forming direct H-bonds (Figure 5).

The results of biological testing of the peptides is demonstrated in Figures 4, 6-10. In co-cultures of PC12-E2 or CGN cells with fibroblasts expressing NCAM the P1-B, P1-CD, P2-A'B, P3-DE, and P3-G peptides all inhibited NCAM-stimulated neurite outgrowth, indicating an impaired NCAM-NCAM binding between the two cell layers. The corresponding control peptides have little or no inhibitory effect (Figure 6G). In contrary, the peptides P2-EF, P2-CD and P3-G did not affect NCAM-stimulated neurite outgrowth in co-cultures and are very capable stimulators of neurite outgrowth in primary cultures of single CGN (Figures 8-10). The peptides P2-A'B, and P1-CD (Figures 8-12) were capable modulating the NCAM homophylic adhesion- mediated neurite outgrowth, but did not stimulate differentiation of single CGN in primary culture.

The P1-B peptide interferes with the Ig1-to-Ig2 interaction and thereby inhibits the Ig1-Ig2-mediated *cis* dimerization of NCAM. In the crystals of the Ig1-2-3 module zipper-like arrays of NCAM *cis* dimers are observed, reflecting *trans* interactions of NCAM. *Trans* interactions therefore seem to require *cis* dimerization of NCAM molecules (Figure 4). The P3-DE and P3-G peptides will not affect *cis* interactions but interfere with *trans* interactions. Since the NCAM-dependent neurite outgrowth relies on NCAM-NCAM interactions between the two cell layers, an inhibition of these interactions will directly affect NCAM-mediated neurite outgrowth.

In our study, we show that mutations in the peptides derived from the Ig3 module produce the same effect as that of the similar mutations in the Ig3 module. This demonstrates that in this experimental setup the employed peptides mimic the Ig3 module, and thus can be used as a convenient and simple tool for further analysis. Moreover, the peptides representing the sequence of the Ig3 module homophilic binding site of chicken NCAM (Ig1-to-Ig3 binding site in the present work) have been

previously used to identify and characterize the Ig3 module homophilic binding site (Rao et al., 1992; Sandig et al., 1994; Rao et al., 1994). These results, combined with the Ig3 mutation studies, provide strong evidence for a biological role of the observed Ig1-to-Ig2, Ig1-to-Ig3, and Ig2-to-Ig3 contacts.

Novel zipper mechanism for NCAM homophilic adhesion

The crystal structure of the Ig1-2-3 fragment reveals novel interactions between the Ig1 and Ig3 and the Ig2 and Ig3 modules of NCAM, as well as shows previously observed Ig1-to-Ig2 and Ig2-to-Ig2 interactions (Kasper et al., 2000). Together, these contacts mediate formation of two perpendicular zipper-like arrays of the Ig1-2-3 dimers (Figure 4). The parallel interaction of the NCAM Ig1-2-3 molecules in the crystal mediated by the Ig1-to-Ig2 contact may reflect an interaction between NCAM molecules present on the same cell surface – *cis* interaction. The anti-parallel interactions mediated by the Ig1-to-Ig3, Ig2-to-Ig2, and the Ig2-to-Ig3 contacts may reflect the interaction of NCAM molecules present on opposing cells - *trans* interactions. Based on all presented observations, we propose a model for NCAM homophilic adhesion, consisting of two zipper-like arrays of NCAM molecules (Figure 7). In the "compact" zipper (Figure 7A), NCAM *cis* dimers originating from opposing cell membranes are arranged as arrays through Ig1-to-Ig3 and Ig2-to-Ig2 interactions. We speculate that "compact" zippers are likely to form first as they allow larger distances between opposing cell membranes than the perpendicular "flat" zippers. In the"flat" zipper (Figure 7B), the Ig2-to-Ig3 interactions suggest a lateral association between the NCAM "compact" zippers thereby forming a double zipper adhesion complex (Figure 7C). The glycosylation at Asn203 of Ig3 (Figure 2) is not likely to interfere with the ability to form the zippers as supported by the fact that the glycosylated Ig3 module inhibits NCAM-mediated neurite outgrowth, whereas glycosylated Ig3mut2 containing mutations at the Ig2-Ig3 binding site is inactive (Figure 6F,G). In the "compact" zipper, the heparin binding sites (133-KHKGRDVILKKDVRFI-148) (SEQ ID NO: 39) (Cole and Akeson, 1989) of Ig1-2-3 molecules are solvent exposed (Figure 2C,D) and therefore accessible for binding to heparin and heparan sulfate molecules, suggesting that NCAM can be engaged in homophilic and heterophilic interactions simultaneously.

In order to accommodate all seven extracellular modules of NCAM within a typical distance between plasma membranes of ~30 nm (Hall and Rutishauser, 1987), a bend has to be introduced in the NCAM molecules in our model (Figure 7). Analyses

of NCAM by electron microscopy have revealed such a bent rod-like structure (Hall and Rutishauser, 1987; Becker et al., 1989). The angle of the bend at the hinge-region between N-terminal (~18 nm) and C-terminal (~10 nm) parts varies considerably (50-140°) with an average value of 98° (Becker et al., 1989) and presumably provides sufficient internal flexibility for NCAM to fit within the cell-cell distance. Based on these studies and on an average length of ~4.3 nm for an Ig module (present work) and ~3.5 nm for a FnIII module (Leahy et al., 1996), the hinge region is most likely located after Ig4. A multiple sequence alignment of NCAM sequences from various species of vertebrates reveals conserved Pro, Lys, and Gly residues in the PKLQGP sequence connecting the Ig4 and Ig5 modules. Since Pro and Gly are typically associated with polypeptide bends, this sequence is likely to introduce a bend between Ig4 and Ig5 modules. The double zipper observed in the crystal (Figure 7C) presents Ig modules 1 to 3 at differing heights, implying that the NCAM molecules upon co-existence of the zippers are bent with different angles. This is in accordance with the electron microscopy data (Hall and Rutishauser, 1987; Becker et al., 1989).

Although *cis* interactions between the Ig1-Ig2 modules do not mediate cell-cell interactions themselves, they probably contribute to the stability of the *trans* interactions. This contention is supported by the cell co-culture experiments using the P1-B peptide corresponding to the site in Ig1 binding to Ig2 (Figure 6). Furthermore, an inhibitory effect on cell aggregation was recently demonstrated for a peptide 172-GRILARGEINFK-182 (P2 peptide) (SEQ ID NO: 28) representing the site in the Ig2 module binding to the Ig1 module (Soroka et al., 2002). Therefore, we suggest that the formation of *cis* dimers may be a prerequisite for the establishment of *trans* interactions.

To our knowledge, only three X-ray structures of Ig module containing adhesion molecules have been determined comprising three or more Ig modules (axonin-1/TAG1 (Freigang et al., 2000), hemolin (Su et al., 1998), and CD4 (Wu et al., 1997). A similar zipper-like array of *trans*-interacting *cis* homodimers has been observed in the crystal structure of the junctional adhesion molecule (JAM) (Kostrewa et al., 2001). A zipper-like mechanism of homophilic interactions was also suggested for axonin-1/TAG-1 (Freigang et al., 2000), where molecules alternately provided by opposed membranes form a linear zipper-like array. However, the double zipper formed by NCAM differs fundamentally from the previously described zippers.

In conclusion, we here present a novel model for NCAM homophilic binding, which is based on the formation of zippers. The model is in agreement with a number of studies demonstrating that the Ig1, Ig2, and Ig3 modules all are involved in NCAM homophilic binding (Rao et al., 1992; Sandig et al., 1994; Kiselyov et al., 1997; Jensen et al., 1999; Kasper et al., 2000; Atkins et al., 2001) and reconciles a large body of conflicting biological data. The crystal structure of the Ig1-2-3 fragment reveals details of two so far unknown interactions between Ig1 and Ig3 and between Ig2 and Ig3. Interestingly, the Ig1 and Ig2 modules of NCAM mediate both *cis* and *trans* interactions simultaneously, whereas Ig3 is involved only in *trans* interactions. All taken together, our study implies that it is the joined forces of the first three Ig modules that confer the strength of the NCAM-mediated adhesion.